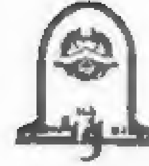
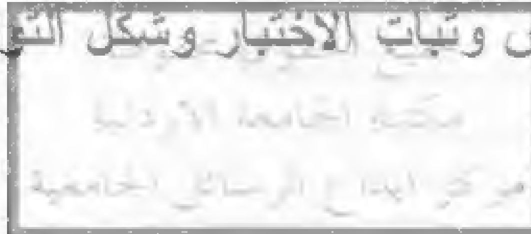
بسم الله الرحمن الرحيم



جـامعة مؤتــــة
عمادة الدراسات العليا

أثر طرق التصحيح لأثر التخمين على الخطأ المعياري للقياس وثبات الاختبار وشكل التوزيع .



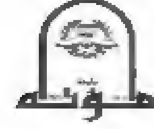
طه عقلة الخرشة

رسـالـة
مقدمــــة إلـــــى
عمــــادة الدراسات العليا
استكمالاً لمتطلبات الحصول على درجة
الماجستير في القياس والتقويم قسم علم النفس

جامعة مؤتة، 2004

بسم الله الرحمن الرحيم

جامعة مؤته



إجازة رسائل جامعية

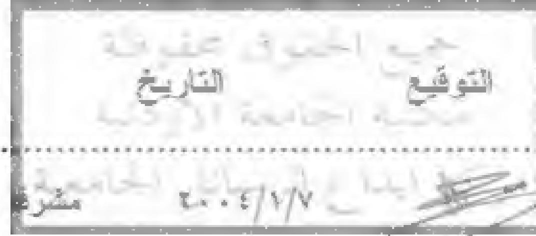
عمادة الدراسات العليا

تقرر إجازة الرسالة المقدمة من الطالب طه عقلة الخرشة والموسومة بـــ:

"اثر طرق التصحيح لأثر التخمين على الخطأ المعياري للقياس وثبات الاختبار وشكل التوزيع".

استكمالا لمتطلبات الحصول على درجة الماجستير في القياس والتقويم

القسم : علم النفس



الاسم	التوقيع	التاريخ	
د. عبد الله الصمادي		٢٠٠٤/١/٧	مشرفا
د. موسى النبهان		٢٠٠٤/١/٧	عضوا
د. عماد الزغول		٢٠٠٤/١/٧	عضوا

عميد الدراسات العليا

د.ذياب البداينـــــة



الإهداء

إلى والدي، إلى والدتي، اعترافا بفضلهما مبتهلا إلى الله أن يجزيهم خير الجزاء.إلى إخواني، وأخواتي، الذين شاركوني أفكارهم. إلى كل من علمني حرفا منذ الصغر كل الحب والتقدير. إلى كل من قدم لي مساعدة في هذه الرسالة. اهدي لهم جميعا هذا الجهد المتواضع .

طه عقلة الخرشة

شكر وتقدير

اعترافا بالجميل وقد أنهيت رسالتي هذه أتقدم بالشكر والعرفان والتقدير لأساتذتي الأفاضل الذين أناروا لي الطريق لتحقيق غايتي بما أسدوه لي من نصح وإرشاد وتوجيه وساعدوني لإخراج هذه الدراسة إلى حيز الوجود .وأخص بالشكر الدكتور الفاضل عبدالله الصمادي المشرف على هذه الرسالة والذي تابع باهتمام مراحل اعدادها خطوة خطوة حتى أنجزت. فالله أسأل أن يجزيه عني خير الجزاء ويحفظه ذخراً وسنداً لأهلة ووطنه وطلابه.

وكل الشكر والتقدير لأعضاء لجنة المناقشة المتمثلة بالدكتور موسى النبهان والدكتور عماد الزغول لتفضلهما بمناقشة هذه الرسالة، وتقديم الملاحظات القيمة والتي أثرت هذه الرسالة، فوجود اسميهما عليها وسام شرف أعتز به وأفخر.

كما أتقدم بجزيل الشكر وعظيم الامتنان إلى أستاذي الفاضل الدكتور ساري سواقد الذي لم يبخل علي بإسداء نصائحه التربوية وتوجيهاته القيمة.

كما وأتوجه بالشكر إلى كل من: الدكتور طه الخرشه و الزملاء مالك الكساسبه وعاطف الصرايرة، وإلى كل من قدم لي يد العون والمساعدة لهم جميعاً كل الشكر والتقدير.

والله ولي التوفيق

طه عقلة الخرشة

فهرس المحتويات

المحتويات — الصفحة

الفصل الثالث: المنهجية والإجراءات

الفصل الرابع : عرض النتائج

الفصل الخامس : مناقشة النتائج والتوصيات

قائمة المراجع

فهرس الجداول

فهرس الأشكال

الصفحة	موضوع الشكل	رقم الشكل
12	اجراءات تقدير معامل الاستقرار..........................	1
13	اجراءات تقدير معامل الاتساق الداخلي	2
13	اجراءات تقدير قيمة معامل التكافؤ..........................	3
16	زيادة قيمة الثبات بزيادة القدر النسبي لتباين العلامات ...	4
52	المدرج التكراري لعلامات النموذج الأول..................	5
52	المدرج التكراري لعلامات النموذج الثاني	6
52	المدرج التكراري لعلامات النموذج الثالث	7

فهرس الملاحق

رقم الملحق	موضوع الملحق	رقم الصفحة
1	الشكل الأولي للاختبار التحصيلي	64
2	الشكل النهائي للاختبار التحصيلي	75
3	الفقرات المحذوفة	82

الملخص

أثر طرق التصحيح لأثر التخمين على الخطأ المعياري للقياس وثبات الاختبار وشكل التوزيع

طـــه عقلة الخرشـــة

جامعة مؤتة،2004

سعت هذه الدراسة إلى المقارنة بين ثلاث طرق لتصحيح أثر التخمين على الخطأ المعياري للقياس وثبات الاختبار وشكل توزيع العلامات، وهذه الطرق هي:

1. معادلة التصحيح لأثر التخمين (العقاب)
2. معادلة التصحيح لأثر التخمين (المكافأة)
3. معادلة التصحيح لأثر التخمين (العقاب والمكافأة) .

تمّ إعداد اختبار تحصيلي يشتمل على(30) فقرة من نوع الاختيار من مـــتعدد ذات أربع بدائل لقياس المصطلحات والمفاهيم الأساسية في مقرر مـــبادئ القياس والتقويم الذي يدرس لطلاب كلية التربية بجامعة مؤتة. وتم تقســـيم فقرات الاختبار إلى ثلاث مجموعات، بحيث اشتملت كل مجموعة منها على(10) فقرات، ورتبت أسئلة هذه المجموعات الثلاث بطرق مختلفة فـــي ثلاث نماذج،أعطيت لعينة مكونة من (197) طالبا وطالبة، صحح كل نموذج منها بإحدى طرق التصحيح الثلاث .

أظهرت نتائج الدراسة، وجود فروق ذات دلالة إحصائية، عند مستوى دلالـــة($\alpha \geq .05$) بيـــن معاملات الثبات، عند استخدام معادلة التصحيح لأثر التخميـــن (العقاب والمكافأة)، بالمقارنة مع استخدام معادلة التصحيح لأثر

التحمين (العقاب) ولصالح معادلة العقاب والمكافأة، في حين لم تكن الفروق بين معاملات الثبات، ذات دلالة إحصائية، عند مستوى دلالة($\alpha \geq 05$) عندما تم تصحيح الاختبار بكل من معادلة التصحيح لأثر التخمين (المكافأة)، أو معادلة التصحيح لأثر التخمين (العقاب والمكافأة). أضافه إلى ذلك، فقد تم الحصول على نقصان في قيمة الخطأ المعياري للقياس عند استخدام معدلة التصحيح لأثر التخمين (العقاب والمكافأة)، بالمقارنة مع استخدام معادلة التصحيح لأثر التخمين (العقاب) ومعادلة التصحيح لأثر التخمين (المكافأة)، وأظهرت النتائج أيضا وجود فروق ذات دلالة إحصائية عند مستوى دلالة($\alpha \geq 05$) بين الأوساط الحسابية لعلامات المفحوصين تعزى لطريقة التصحيح ولصالح طريقة التصحيح القائمة على المكافأة. كذلك، فقد تم الحصول على اختلاف في شكل توزيع العلامات تبعاً لطريقة التصحيح، فقد كان شكل توزيع العلامات أقرب إلى التوزيع الطبيعي عند استخدام معادلة التصحيح لأثر التخمين (العقاب والمكافأة)، بالمقارنة مع استخدام معادلة التصحيح لأثر التخمين (العقاب) ومعادلة التصحيح لأثر التخمين (المكافأة)

وعلية فان نتائج هذه الدراسة توصي باستخدام معادلة التصحيح لأثر التخمين (العقاب والمكافأة) لتصحيح اختبارات الاختيار من متعدد. وذلك لانها تعطي اكبر تقدير لمعامل الثبات.

Abstract

The Effect of the Correction-for-Guessing Formulae on the Reliability of Tests, Standard Error of Measurement and Distribution's Characteristics.

Taha Okleh Al- Khursheh
Mu'tah University,2004

This study sought for comparing between three formulae for correcting the effect of guessing on the reliability of tests, standard error of measurement and distribution's characteristics. These formulae are the punishment, rewarding, and the punishment and rewarding Formula

An achievement test was prepared. It consisted of thirty multiple choice items with four alternative answers for each. It measures the students' attainment of the basic concepts of "the Principles of Measurement and Evaluation Course" in the Faculty of Education at Mu'tah University

The test items were divided into three groups. Each group included ten items, and the questions were differently arranged in three models. The models were given to a sample of 197 students, and then each model was graded using one of the three correction formulae.

The results of the study uncovered a statistically important differences at the significant level $(05 \geq \alpha)$of the reliability coefficients when using the Punishment and Rewarding Formula compared with the Punishment or Rewarding Formula. In addition, the standard error of measurement was reduced when the Punishment and Rewarding Formula was used in comparison to using of Punishment or Rewarding Formula. The results also showed the existence of statistically important differences in favor of the Punishment Formula at the significant level $(05 \geq \alpha)$

between the means of the examinees' grades that are ascribed to the correction formulae.

As a result, the study recommends the usage of the punishment and rewarding Formula to Correction-for-Guessing when grading multiple choice tests in order to get the highest levels of reliable testing.



الفصل الأول

خلفية الدراسة وأهميتها

المقدمـــة:

يعد التقويم عنصراً أساسياً من عناصر المنهج التعليمي وهو بذلك أحد مرتكــرات تطوير التعليم، إذ لا يمكن إحداث تطوير في أهداف ومحتوى المـــنهج التعليمي وطرائق التدريس إلا بالاعتماد على نتائج التقويم. ومن هنا فإنه لابد أن ينظر لتقويم الطالب نظرة اهتمام خاصة من حيث التخطيط له وإعداد أدواته والاستفادة من نتائجه لكي يتمكن المعلم من التعرف على مـــدى تحقق الأهداف التي يسعى إلى تحقيقها. وإذا كان تقويم النتائج في عمـــل المهندس والط[illegible]ح والفاعليه، فإنه فـــي عمل المعلم لا [illegible]ي الأسمى وتأثيره هـــو الأبلـــغ والأجـ[illegible]توقعات ويحدد القيم والتوجهات ويرسم إطار مستقبل الأمة.

تعد اختبارات الاختيار من متعدد من أهم أدوات تقويم الطالب وأكثرها شيوعاً واستخداما، وربما يعود ذلك لما يمتاز به هذا النوع من الاختبارات مـــن قـــدرة على تمثيل المجال السلوكي للسمة موضوع القياس، وسهولة التصحيح، والموضوعية . ..، غير أن هذه الاختبارات، لا تخلو من أخطاء القـــياس، والتـــي تؤثر بشكل كبير على عملية التقييم والقرارات المترتبة علـــيها والتي تكون أحيانا غير صحيحة نتيجة هذه الأخطاء (Ban-s.mon & Budescu,1997).

وقـــد يتســـاءل البعض عن مصادر هذه الأخطاء من أجل التحكم فيها أثناء إعداد الاختبار وعند تطبيقه وتصحيح فقراته، وذلك لجعل تأثيرها في

العلامات الحقيقية اقل ما يمكن، مما يؤدي إلى ان تقترب العلامات الملاحظة للإفراد على الاختبار من علاماتهم الحقيقية التي تعبر عن مدى امتلاكهم للسمة موضوع القياس، وبالتالي تزداد دقة القرارات المترتبة على ذلك(علام، 2002).

ومن أهم هذه الأخطاء التخمين العشوائي الذي يشير إلى السلوك الذي يقوم به المفحوص عندما لا تكون لديه الإجابة الصحيحة؛ أي عندما لا تكون لديه المعرفة التامة والأكيدة بالإجابة الصحيحة للفقرة. ففي مثل هذه الحالة؛ إما أن تكون لدى المفحوص معرفة جزئية(غير تامة أو أكيدة) "Partial Knowledge" تجعله قادرا على استبعاد بعض البدائل وحصر الإجابة الصحيحة في البعض المتبقي من البدائل الموضوعة للفقرة، فيختار أحدها كإجابة للفقرة، وقد لا تكون لديه مثل هذه المعرفة الجزئية فيختار أحد البدائل عشوائيا كإجابة للفقرة(سواقد، 1992).

وتعد مشكلة التخمين في اختبارات الاختيار من متعدد من بين المشكلات التي اهتم خبراء القياس النفسي والتربوي بدراستها وبحثها للتعرف على مدى تأثيرها سواء في أداء المفحوص أو في مختلف الخصائص السيكومترية لهذه الاختبارات(علام،1990). و يعود ذلك إلى كون التخمين يمثل أحد مصادر الخطأ في القياس، وعليه فقد تمّ اقتراح عدد من الطرق للتغلب على هذه المشكلة والتقليل من تأثيرها، وأكثر هذه الطرق شيوعا، طريقة عدد الإجابات الصحيحة "Number - right scoring" وطريقة التصحيح لأثر التخمين(العقاب) وطريقة التصحيح لأثر التخمين (المكافأة)(Abu-sayf,1977).

وجميع هذه الطرق تقوم على مجموعة من الافتراضات فأولها: أن المفحوص يعرف الإجابة الصحيحة، وثانيها: أن المفحوص لا يعرف

الإجابــة الصــحيحة ويــترك الفقرة دون إجابة، وثالثها: أن المفحوص لا يعرف الإجابة على الفقرة، ولكنه يجيب على الفقرة بشكل عشوائي (Ben-shakhar&Sinai,1991).

وقــد لوحــظ أنّ بعــض المستخدمين لاختبارات الاختيار من متعدد يســتخدمون طــريقة التصــحيح التقليدية، ويستخرجون نتائج يبنون عليها قــرارات وتوصــيات متعددة، رغم أن جميع الدراسات في هذا الموضوع أثبتــت عدم كفاية هذه الطريقة لتصحيح اختبارات الاختيار من متعدد من حيث قدرتها على ضبط عامل التخمين.

ولذلــك تأتــي هذه الدراسة للكشف عن طريقة التصحيح التي تجعل علامــة المفحــوص على الاختبار أكثر دقة في تمثيل قدرته الحقيقية على السمة موضوع القياس، وذلك في ضوء قدرتها على ضبط التخمين.

مشكلة الدراســة:

من ابرز المشاكل التي يعاني منها القياس النفسي والتربوي هي وجود خطأ القياس حيث يشكل التخمين أحد مصادر الخطأ الأساسية في القياس. إذ أن المفحوص قد يحصل باستخدام التخمين على علامة اكبر من العلامة التي تمثل قدرته الحقيقية على السمة موضوع القياس. والسؤال الذي يطرح نفســه هــو : هل توجد طريقة تصحيح أكثر دقة للفقرات التي يؤثر فيها التخمين بحيث تمثل العلامة على الاختبار تقديراً دقيقاً لقدرة المفحوص في الســمة المقاســة. وهــل طــريقة التصحيح تؤدي إلى تحسين الخصائص السيكومترية للاختبار ؟

وللإجابــة علــى هــذه الســؤال استخدم العاملون في الميدان طرقا ومعالجــات إحصائية اعتمدت جميعها على طريقة تصحيح الاختبار. من

بينها توظيف معادلات خاصة للتصحيح لأثر التخمين منها معادلة عقاب التخمين. إلا أن دراسات عديدة شككت بفعالية استخدام معادلة عقاب التخمين في تحسين ثبات الاختبار. ولما كانت الدراسات التي توفرت لا تشير إلى وجود دراسات تجريبية لاختبار طريقة المكافأة على عدم التخمين أو طريقة المكافأة والعقاب، أتت هذه الدراسة لتقارن بين ثلاث طرق لتصحيح أثر التخمين في اختبارات الاختيار من متعدد على ثبات الاختبار والخطأ المعياري للقياس وشكل التوزيع والطرق الثلاث هي :

1- استخدام معادلة التصحيح لأثر التخمين(العقاب)
2- استخدام معادلة التصحيح لأثر التخمين(المكافأة)
3- استخدام معادلة التصحيح لأثر التخمين(المكافأة والعقاب)

أسئلة الدراسة:

ستحاول هذه الدراسة الإجابة عن الأسئلة الآتية:

1- هل لطريقة التصحيح أثر على ثبات الاختبار؟ أي هل توجد فروق ذات دلالة إحصائية بين معاملات ثبات الاختبار عند تصحيحه بالطرق الثلاث(عقاب التخمين،مكافأة عدم التخمين،العقاب والمكافأة).

2- كيف تؤثر طريقة التصحيح على الخطأ المعياري للقياس. أي هل توجد فروق ذات دلالة إحصائية بين قيمة الخطأ المعياري للقياس لكل طريقة من طرق التصحيح السابقة ؟

3- هل تختلف خصائص توزيع درجات الاختبار باختلاف طريقة التصحيح لأثر التخمين ؟

أهمية الدراســة:

لقــد أشار(عــودة،1993)إلى ان اختبارات الاختيار من متعدد. أصبحت الأكثر انتشارا بين كل أدوات جمع البيانات في التربية والمجالات التنظيمية الأخرى، ومما زاد في شيوع وانتشار هذا النوع من الاختبارات أن فقرات الاختيار من متعدد تفوقت على أشكال الفقرات الموضوعية الأخرى وربما الأســئلة المقالية أيضا، وذلك بسبب كفاءتها وتعدد استعمالاتها، حيث يمكن بواســطتها قــياس أهــداف بسيطة وأخرى مركبة في مختلف المواضيع الدراسية والمراحل التعليمية. واكثر الطرق شيوعا لتصحيح هذا النوع من الاختبارات طريقة التصحيح التقليدية بحيث تمثل علامة المفحوص مجموع علامات الفقرات للمفحوص، وتمنح علامة واحدة للإجابة الصحيحة علــى الفقــرة وصفر لأي إجابة أخرى، كما يعتبر استخدام طريقة عقاب التخمين الأكثر شيوعا في التصحيح بعد الطريقة التقليدية. وتشير الدراسات التجريبية التي أجريت للمقارنة بين طرق تصحيح الاختبارات الموضوعية، إلى عدم قدرة أي من الطريقتين على ضبط عامل التخمين وأثره على ثبات الاختــار. ولــم تتوفر في حدود معرفة الباحث دراسات تجريبية لاختبار الطــريقة الثانــية والثالــثة. لذا تأتي أهمية هذه الدراسة لمحاولة تجريب الطريقة الثانية والثالثة ومقارنة نتائجهما مع نتائج الطريقة الأولى، وكذلك للكشــف عن طريقة التصحيح التي تجعل علامة المفحوص على الاختبار اكــثر دقة في تمثيل قدرته الحقيقية على السمة موضوع القياس في ضوء قدرتها على الحد من استخدام المفحوص للتخمين العشوائي، وانعكاس ذلك فــي اثر هذه الطرق على كل من الخطأ المعياري للقياس وثبات الاختبار وشكل توزيع العلامات .

متغيرات الدراســة:

المتغــير المستقل في هذه الدراسة، طريقة تصحيح الاختبار المتمثلة بالطرق الثلاث التالية :-

1- استخدام معادلة التصحيح لأثر التخمين(العقاب)

2- استخدام معادلة التصحيح لأثر التخمين(المكافأة)

3- استخدام معادلة التصحيح لأثر التخمين(المكافأة والعقاب)

إما المتغيرات التابعة فهي :-

معامل ثبات الاختبار ويحسب لعلامات المفحوص على الاختبار باستخدام معادلة كرونباخ الفا(α)

قيمة الخطأ المعياري للقياس وتحسب قيمة هذا الخطأ باستخدام الصيغة الرياضية رقم (١)

التعريفات الإجرائية:

1- طــريقة التصحيح : تحديد القاعدة التي تعطى على أساسها قيم عددية لفقرات الاختبار؛ أي أنها طريقة لتحديد مقياس معين للمعلومات التــي نحصل عليها من إجراءات القياس، يعكس مستوى الأداء عند المفحوصين على الاختبار(سواقد،1992).

الطرق الثلاث لتصحيح الاختبار:

أ- طريقة تصحيح الاختبار من نوع الاختيار من متعدد باستخدام معادلة التصــحيح لأثر التخمين(العقاب) وتكون علامة المفحوص على الاختبار المصحح بهذه الطريقة تساوي :

$$\text{عدد الإجابات الصحيحة} - \frac{\text{عدد الإجابات الخاطئة}}{\text{عدد البدائل} - 1}$$

ب- طريقة تصحيح الاختبار من نوع الاختيار من متعدد باستخدام معادلة التصحيح لأثر التخمين(المكافأة) وتكون علامة المفحوص على الاختبار المصحح بهذه الطريقة تساوي :

$$\text{عدد الإجابات الصحيحة} + \frac{\text{عدد الفقرات المحذوفة}}{\text{عدد البدائل}}$$

جـ- طريقة تصحيح الاختبار من نوع الاختيار من متعدد باستخدام معادلة التصحيح لأثر التخمين(المكافأة والعقاب) وتكون علامة المفحوص على الاختبار المصحح بهذه الطريقة تساوي :

$$\text{عدد الإجابات الصحيحة} + \frac{\text{عدد الفقرات المحذوفة}}{\text{عدد البدائل}} - \frac{\text{عدد الإجابات الخاطئة}}{\text{عدد البدائل} - 1}$$

2 ثبات الاختبار : هو معامل الارتباط بين مجموعة العلامات التي حصل عليها المفحوصين في اختبار ما ومجموعة علامات حصل عليها نفس المفحوصين في اختبار مكافئ ([illegible]).

3- الخطأ المعياري للقياس : حاصل ضرب الانحراف المعياري للدرجات الملاحظة في الجذر التربيعي لمعامل الثبات مطروحا منه الواحد الصحيح(علام،2002).

محددات الدراســة:

هــناك بعض المحددات والتي ينبغي تذكرها لدى تعميم ما تسفر عنه الدراسة من نتائج وهي:

- افــتراض أن جمــيع المفحوصين قد فهموا التعليمات المعطاة لهم لأداء الاختبار.

صغر حجم العينة، إذا اقتصرت عينة الدراسة التي جرى تحليل نتائجها على (197) مفحوصا.

الفصل الثاني
الإطار النظري والدّراسات السّابقـة

الإطار النظري:

تشـكل علـوم القياس جانباً مهما في دراسة الظواهر المختلفة، فهي تختـص بقياس الظاهرة وتقديرها. وكلما ازداد القياس موضوعية تحسنت درجـة فهمنا للظاهرة، وأدى هذا إلى دقة في التنبؤ وما يتبع ذلك من دقة في الضبط والتحكم(الشرقاوي والشيخ وكاظم وعبد السلام، 1996).

فالغرض الأساسي من القياس هو جمع المعلومات الصادقة لكي تساعد في اتخاذ قرارات تتعلق بظاهرة محددة وتعتمد صحة القرار إلى حد كبير علـى مـدى شمولية ودقة المعلومات التي جمعت. وإلا فان الخطأ يزداد باتسـاع من يشملهم هذا القرار(الثبيتـي، 1998). وتستخدم أدوات القياس في جمع هذه المعلومات بطريقة منظمة تفيد في صنع القرارات المختلفة، وهذه الأدوات تشـتمل علـى مجموعة من المطالب أو المهام أو المثيرات التي ترتـبط بسمة نفسية أو تربوية، وعادة تتطلب أدوات القياس من الفرد أن يستجيب لهـذه المطالـب (فقرات أو أسئلة) بحيث يمكن الاستدلال من اسـتجابته علـى مقدار السمة المراد قياسها(عـلام،2002). ومن بين أدوات القياس الاختبارات.

الاختبارات:

إن تعـدد السمات واختلاف طبيعتها وتباين أنواعها يتطلب أن تتباين أدوات قياس هذه السمات، لذلك لا نجد تعريفاً مناسباً ينطبق على جميع هذه الأدوات. فمـثلا يعـرف ساكس الاختبار النفسي والتربوي بأنه مطلب أو

مجموعـــة مـــن المطالب تستخدم للحصول على ملاحظات منظمة يفترض أنهـــا تمـــثل ســـمة نفســـية أو تـــربوية(عـــلام،2002). ويعـــرّفه كرونباخ (Cronbach,1984) بأنـــه " طريقة منظمة لمقارنة سلوك شخصية بسلوك شخصـــية أخـــرى "والمقصود بمفهوم السلوك في علم النفس أي تصرف صادر عن الفرد كاستجابة لوجود منبه أو مثير، و أما التنظيم فهو خاصية مـــن خصـــائص العلـــم، وهـــذا يشير إلى أن الاختبار في نظر كرونباخ "Cronbach" هو أحد الوسائل العلمية لجمع المعلومات عن جوانب محددة في الشخصية؛ أي أنه أحد وسائل القياس الكمية.

أما عودة (2000) فيعرف الاختبار بشكل عام بأنه أداة قياس يتم أعداده وفق طريقة منظمة من عدة خطوات تتضمن مجموعة من الإجراءات التي تخضـــع لشروط وفق قواعد محددة، وتهدف لتحديد درجة امتلاك الفرد لسمة أو قدرة معينة من خلال إجاباته عن عينة من الفقرات التي تمثل السمة أو القـــدرة المـــرغوب قياسها. وهذا يتفق مع تعريف براون(Brown,1983) للاختبار بأنه إجراء منظم لقياس عينه من السلوك.

استخدام الاختبارات:

تســـتخدم الاختـــبـــارات في قياس السمات وتحديد مقاديرها لأغراض مخـــتلفة: فقد تستخدم في مسح الواقع أو التنبؤ بما يمكن ان يحدث لظاهرة مـــا أو تحديـــد نواحي القوة والضعف في الظاهرة التي نقيسها. وتستخدم الاختـــبارات فـــي كل الميادين والمجالات. ففي التربية يستخدم المعلمون الاختـــبـــارات للكشـــف عـــن قدرات طلابهم وقياس مستواهم التحصيلي و الـــتعرف على مشكلاتهم ونواحي القوة والضعف عندهم، كما تستخدم في غايـــات تصنيف الطلاب، وقياس ذكائهم وميولهم. وفي علم النفس تستخدم

الاختــدرات فـــي قياس قدرات الإنسان والتعرف على شخصيته والعوامل التي تؤثر في سلوكه(عبيدات وعباس وعبد الحق، 1982).

أي أن الاختــبار وســيلة تمكن القائمين على التعليم من التعرف على مـــدى تحقق الأهداف التعليمية، كما أنه أداة رئيسة من أدوات القياس وهو بذلـــك ينبغي أن يكون دقيقاً خالياً من الخطأ، وبما أن الاختبار هو أداة من أدوات القـــياس، لذا ينبغي أن تتوفر فيه بعض الصفات الأساسية كي نثق بنتائجه ونستفيد منها، ومن هذه الصفات:

أولاً : الموضوعية(Objectivity)

و تعني عدم توقف علامة المفحوص على من يصحح ورقته، أو عدم اخـــتلاف علامته باختلاف المصححين، كما تعني أيضا أن يكون الجواب محـــددا ســـلفا بحيث لا يختلف عليه اثنان كما هو الحال في الأسئلة الموضوعية، والموضوعية صفة أساسية من صفات الاختبار الجيد يتوقف علـــيها ثـــبات الاختـــبار وصدقه. وتشتد الحاجة إلى التأكد من موضوعية التصـــحيح فـــي الاختـــبارات المقالية ذلك انها اكثر عرضة للتأثر بذاتية المصحح(الحسن،2000).

ثانياً: الصدق (Validity)

يقصـــد بصدق الاختبار أن ينجح الاختبار في قياس ما وضع لقياسه وتختلف الاختبارات عموما في مستويات صدقها تبعا لاقترابها أو ابتعادها مـــن تقديـــر تلك الصفة التي تهدف لقياسها، والصدق له أهمية كبيرة في تحديد قيمة الاختبار ومغزاه(الشرقاوي والشيخ وكاظم وعبد السلام، 1996).

أمـــا ثورندايك(Thorndike,1977) فيعرف الصدق من خلال التساؤل التالـــي:"هل يقيس الاختبار ما يدعي أنه يقيسه، أم لا ". أما لمهن ومهران (Lehmann & Mehrens,1978) فيؤكدان على أن خاصية الصدق هي أهم

جانـــب فـــي الاختبارات، وعلى أن أفضل تعريف للصدق هو:مدى قدرة الاختبار على تحقيق أهداف معينة .

ثالثاً : الثبات (Reliability)

ثبات الاختبار هو خاصية أخرى يجب التحقق منها للتأكد من صلاحية الاختبـــار قبل تطبيقه، ويقصد بالثبات مدى دقة قياس الاختبار للصفة التي يقيســها، وبعبارة أخرى يقال إن الاختبار ثابت إذا قاس الصفة بدقة وثبات مـــن مرة إلى أخرى وفي ظروف مماثلة، كأن يطبق اختبار تحصيلي في سنة ما في مادة العلوم لمستوى صفي معين ويعاد تطبيق نفس الاختبار في الســـنة التـــي تلي ويتم إيجاد معامل الارتباط بين الاختبارين لمعرفة مدى ثبات أداء الشخص.

ويعتـــبر حرو[illegible] مفهوم [illegible]ات يشير إلى النتائج المعطـــاة مـــن خـ[illegible] قياس ما ونفس الأداة نفسها، أما عباس(1993) فيعرف الثبات بأنه " مدى التغير الناتج عن الأخطاء العشوائية في درجات شـــخص ما عند تطبيق عدة اختبارات متوازية عليه، أو عند تطبيق الأداة نفسها مرات عديدة.

ويعـــرف كل من كروكر و الجينا(Crocker & Algina,1986) الثبات إلـــى انـــه قدرة الاختبار على إعادة نفس النتائج إذا قدم مرة أخرى لنفس المفحوصين في نفس الظروف أو في ظروف مشابهة.

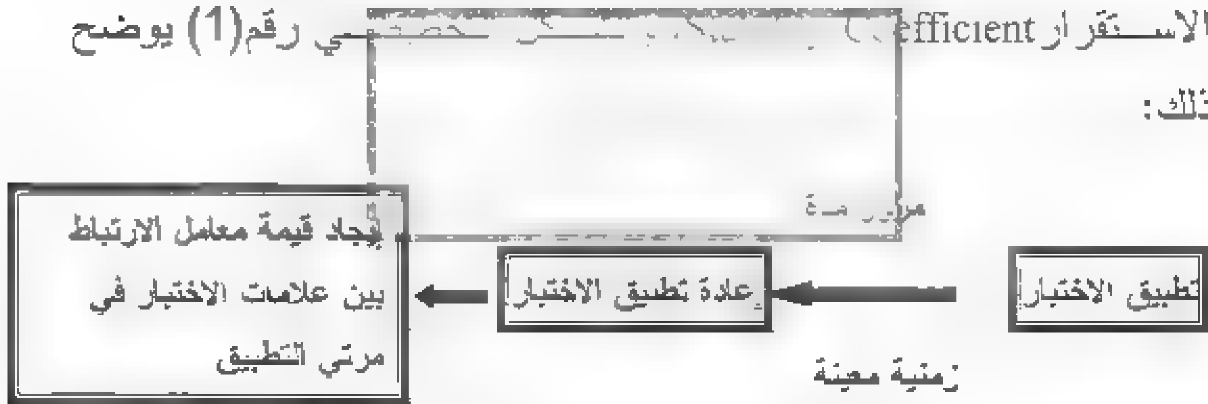
طرق حساب ثبات الاختبار:

للتحقق من الثبات من المفضل أو المرغوب فيه وجود مجموعتين من القياســـات في ظل ظروف متطابقة أو متشابهة، ولكن هذا الإجراء صعب المنال(Grounlund,1985). أما عودة (2000) فيؤكد على وجود أكثر من

طريقة لتقدير معامل ثبات الاختبار وذلك نظراً لتعدد مصادر أخطاء القياس وبالتالي تتأثر الطريقة الواحدة بنوع أو أكثر من أنواع الخطأ، بمعنى أن بعض الأخطاء تظهر في نوع معين من معاملات الثبات ولا تظهر في نوع آخر وهذه الطرق هي :

1- إعادة الاختبار Test - Retest Method

تعتمد هذه الطريقة على إجراء الاختبار على عينة كافية وممثلة، ثم إعادة نفس الاختبار مرة أخرى على نفس المفحوصين بعد فترة زمنية محددة، ويحسب معامل الارتباط بين درجات المفحوصين في المرتين، فيكون معامل الارتباط هذا هو مقياس ثبات الاختبار، ويسمى معامل الاستقرار (efficient [illegible] ي رقم(1) يوضح ذلك:



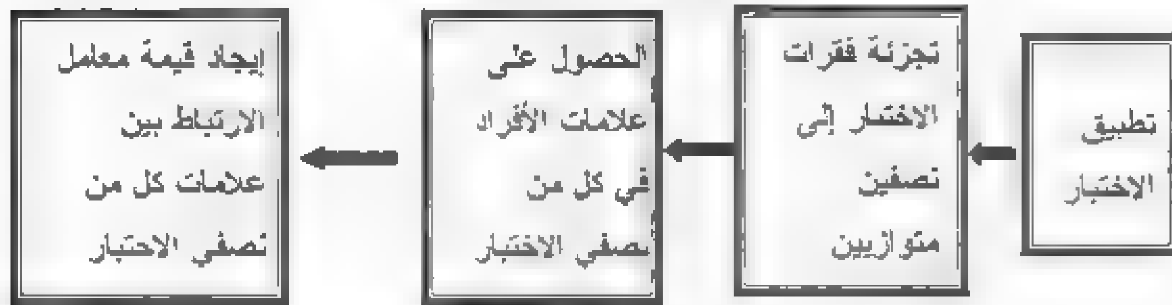
شكل(1)

إجراءات تقدير معامل الاستقرار

2- التجزئة النصفية Split - Half Method

وتعتمد هذه الطريقة على تقسيم الاختبار إلى نصفين بعد تطبيق الاختبار وقبل تصحيحه دون معرفة المفحوص، ثم يضع المصحح درجتين لكل مفحوص، درجة عن النصف الأول ودرجة عن النصف الثاني، ثم يحسب معامل الارتباط بين درجات المفحوصين على نصفي الاختبار، فيكون معامل الارتباط هذا هو مقياس ثبات الاختبار الذي يسمى معامل الاتساق الداخلي "Internal Consistency" (Crocker & Algina,1986)

ويمكـــن توضـــيح إجراءات تقدير قيمة معامل الاتساق الداخلي لاختبار أو مقياس معين باستخدام طريقة التجزئة النصفية بالشكل التخطيطي رقم(2):

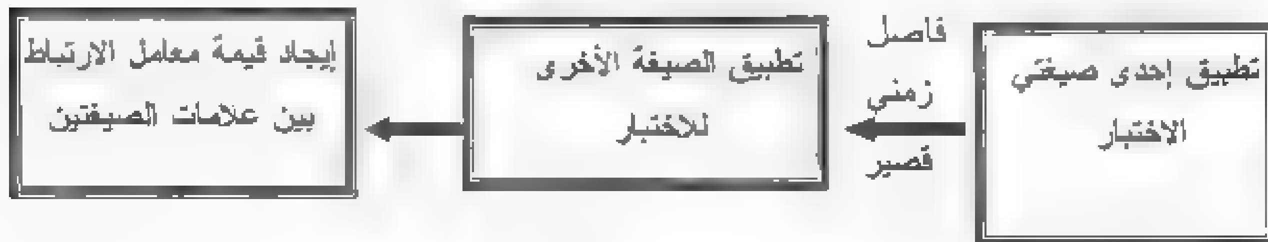


شكل(2)

إجراءات تقدير معامل الاتساق الداخلي

3- الصور المتكافئة Equivalent Forms Method

تستخدم هذه الط[illegible] تماما بحيث تتفق في القيم المختلفة لمق[illegible] لانحراف المعياري ومعـــامل الصدق الداخلي ..الخ. ويتم إيجاد معامل الثبات من خلال إيجاد معـــامل الارتـــباط بيـــن أداء الأفراد على صورتين أو نموذجين متكافئين للاختبار، والذي يسمى معامل التكافؤ "Equivalent Coefficient"(حويج،2002). ويمكن تمثيل إجراءات هذه الطريقة بالشكل التخطيطي رقم (3) التالي:-



شكل(3)

إجراءات تقدير قيمة معامل التكافؤ

الثبات والعلامات الحقيقية:

إن العلامــــة التي يحصل عليها الفرد في اختبار معين تسمى العلامة الملاحظـــة "Obtained Score". غــير أن هذه العلامة تكون في كثير من الأحـــيان مشوبة بأخطاء القياس التي إذا أمكن تحديد مقدارها فإننا نحصل علـــى درجة الخطأ "Error Score". وإذا طرحنا درجة الخطأ من العلامة الملاحظـــة، فإنـــا نحصل على درجة الفرد خالية من الأخطاء العشوائية، وهذه تسمى العلامة الحقيقية"True Score"، أي أن:

العلامة الملاحظة = العلامة الحقيقية + درجة الأخطاء العشوائية

$$X = T + E \ldots\ldots\ldots \ldots\ldots\ldots (1)$$

X : العلامة الملاحظة

T : العلامة الحقيقية

E : الخطأ في القياس

الافتراضات الأساسية لهذه المعادلة هي :

1. الوسط الحسابي للأخطاء العشوائية يساوي صفر،($\bar{E} = Zero$). بمعنى أن قيمة الخطأ تختلف مقداراً واتجاهاً. وعند تكرار القياس العديد من المرات تؤول هذه القيمة إلى الصفر .
2. معامل الارتباط بين العلامات الحقيقية وأخطاء القياس يساوي صفر ($\rho_{TE} = 0$) (Crocker & Algina, 1986).

فالعلامـــة التـــي يحصل عليها الفرد في الاختبار إذن ليست هي العلامة الحقيقية، وإنما هي العلامة التي تتكون من العلامة الحقيقية مضافا إلـــيها درجـــة الأخطـــاء العشـــوائية كما في المعادلة(1)؛ أي أن العلامة الملاحظة للفرد في الاختبار لا تساوي علامته الحقيقية إلا إذا كانت درجة

الخطـــأ صـــفراً، ولكـــنها تزيد عن علامته الحقيقية إذا كانت درجة الخطأ موجبة، وتقل عنها إذا كانت درجة الخطأ سالبة.

فإذا لجأ الفرد إلى التخمين في الإجابة عن فقرات الاختبار وتوصل إلـــى الإجابة الصحيحة فان درجته الملاحظة تزيد عن درجته الحقيقية، إما إذا كـــان الفـــرد متعبا أو أن فقرات الاختبار أو تعليماته غامضة أو نسي سؤالاً فان درجته الملاحظة سوف تقل عن درجته الحقيقية(علام،2002).

المفهوم الإحصائي لمعامل الثبات:

على الرغم من أن العلامات الحقيقية ودرجات الأخطاء العشوائية تعد من المفاهيم الافتراضية [illegible] التوصل إلى المفهوم الإحصـــائي للثبات، [illegible] مرتبطة بالعلامات الحقيقـــية فإن تباين العلامات الملاحظة = تباين العلامات الحقيقية +تباين الخطأ؛أي

$$\sigma_T^2 + \sigma_E^2 = \sigma_X^2 \text{........}(1)$$

بقسمة طرفي المعادلة(1) على σ^2_x نحصل على

$$\frac{\sigma_T^2}{\sigma_X^2} + \frac{\sigma_E^2}{\sigma_X^2} = 1$$

أي أن

$$\frac{\sigma_T^2}{\sigma_X^2} = 1 - \frac{\sigma_E^2}{\sigma_X^2} \text{..............}(2)$$

والقيمة $\frac{\sigma_T^2}{\sigma_X^2}$ تدل على نسبة التباين الحقيقي إلى التباين الملاحظ

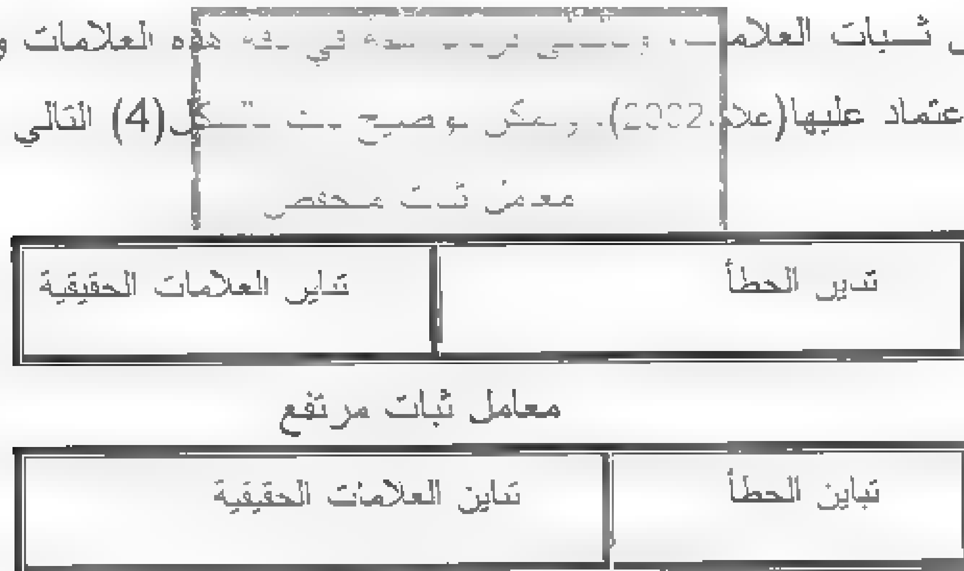
وهـــذا مـــا يعـــبر عـــنه بعـــامل ثبات درجات الاختبار Coefficient of Reliability

إما $\frac{\sigma_E^2}{\sigma_X^2}$ فتدل على نسبة تباين الخطأ إلى التباين الملاحظ

وبالتالي فإن معامل الثبات من المنظور الإحصائي يمكن اعتباره نسبة تباين العلامات الحقيقية إلى تباين العلامات الملاحظة كالتالي :

$$\text{معامل الثبات} = \frac{\text{تباين العلامات الحقيقية}}{\text{تباين العلامات الملاحظة}}$$

وهـذا يعنـي أن معـامل الثبات هو ذلك الجزء من تباين العلامات الملاحظــة الــذي يعــزى إلى تباين العلامات الحقيقية. وتتراوح قيم هذا المعــامل بيــن صفر،1، ولذلك يعبر عن قيمته بكسر عشري مثل معامل الارتــباط. فكلما اقتربت قيمته من الواحد الصحيح دل ذلك على قدر أكبر مــن ثــبات العلامــات، وبالتالي نزداد ثقة في دقة هذه العلامات واتساقها والاعتماد عليها(علام،2002). ويمكن توضيح ذلك بالشكل(4) التالي :



شكل(4)

زيادة قيمة الثبات بزيادة القدر النسبي لتباين العلامات الحقيقية

العلاقة بين العلامات الحقيقية والعلامات الملاحظة(مؤشر الثبات):

تعرفـنا فـيما سبق على مفاهيم العلامة الحقيقية وخطأ القياس، ولقد أصـبح مـن الواضح انه عندما يعطي مدرسٌ أو باحثٌ أو طبيبٌ اختباراً فـإنهم يعـرفون العلامــات الملاحظة فقط، على الرغم من انهم معنيون

بالعلامـــة الحقيقـــية، وهـــنا يبرز السؤال حول مدى التقارب بين العلامة الحقيقية والملاحظة للمفحوص؟! وأحد المؤشرات لهذه العلاقة هو الارتباط بين هذه المتغيرات. ويعرف معامل الارتباط الذي يبين العلاقة بين العلامة الحقيقـــية والعلامـــة الملاحظة في الاختبار بـــ مؤشر الثبات، وعليه فإن العلامة الملاحظة للمفحوص يمكن أن تكون

$$X = T + E$$

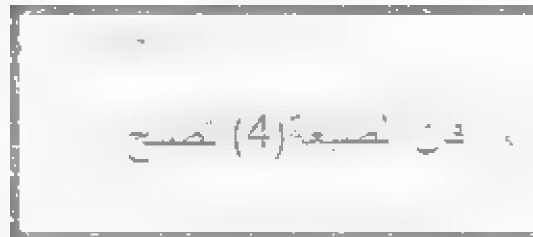
وبالعلامة المعيارية

$$x = t + e \quad \ldots \quad (3)$$

ولإيجاد معامل الارتباط نستخدم الصيغة التالية :

$$\rho_{XT} = \frac{\sum xt}{N\sigma_X\sigma_T}$$

وبتعويض قيمة



$$\rho_{XT} = \frac{\sum (e+t)t}{N\sigma_X\sigma_T}$$

$$\rho_{XT} = \frac{\sum t^2}{N\sigma_X\sigma_T} + \frac{\sum te}{N\sigma_X\sigma_T} \quad \ldots\ldots\ldots\ldots\ldots\ldots (5)$$

ولكن الارتباط بين العلامة الحقيقية وخطأ القياس يفترض انه يساوي صفر؛أي $\frac{\sum te}{N\sigma_X\sigma_T} = 0$ وكذلك $\sigma_T^2 = \frac{\sum t^2}{N}$

وبالتعويض في المعادلة 5 نجد أن :

$$\rho_{XT} = \frac{(\sigma_T)^2}{\sigma_X\sigma_T} = \frac{\sigma_T}{\sigma_X} \quad \ldots \quad (6)$$

ولكـــن ســـبق توضـــيح أن : معامل الثبات $\rho = \frac{\sigma_T^2}{\sigma_X^2}$ فبالتعويض في المعادلة 6 نجد أن : $\rho_{XT}^2 = \rho$ ؛ أي أن $\rho_{XT} = \sqrt{\rho}$

و يعنـــي هـــذا أنه يمكن تقدير معامل الارتباط بين العلامات الحقيقية والعلامـــات الملاحظـــة لمجموعة من الأفراد في اختبار ما بإيجاد الجذر

التربيعي لمعامل ثبات علامات هذا الاختبار.ويسمى معامل الارتباط عندئذ مؤشر الثبات"Reliability Index"(Crocker & Algina.1986؛علام،2002).

الخطأ المعياري للقياس:

نلاحــظ مما سبق أن علامة الفرد الملاحظة في الاختبار تختلف في غالــب الأحيان عن درجته الحقيقية نظرا لتأثر العلامة الملاحظة بمصادر أخطاء متعددة. فإذا افترضنا أننا استطعنا تحديد درجة الأخطاء العشوائية التـــي أثــرت في العلامة الملاحظة لكل فرد، فإنه يمكن إيجاد الانحراف المعــياري لدرجات الخطأ، والقيمة الناتجة تسمى الخطأ المعياري للقياس " Standard Error Of Measurement " [illegible]ـطيع في واقع الأمر إيجاد درجة الخطأ لكل فرد من [illegible] مجموعة [illegible] أعيد الاختبار على الفــرد نفســه عدد كبير من المرات، وهو أمر غير ممكن؛ لذلك فإننا لا نستطيع إيجاد الخطأ المعياري للقياس، ولكننا نستطيع تقدير هذه القيمة إذا علمــنا قيمة الانحراف المعياري للعلامات الملاحظة، وكذلك قيمة معامل ثــبات درجــات الاختـــار، وذلك باستخدام صيغة رياضية يمكن اشتقاقها مباشرة من المعادلة رقم(2) كما يلي :

$$\frac{\sigma_E^2}{\sigma_X^2} = 1 - \frac{\sigma_T^2}{\sigma_X^2} \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

وبما أن $\frac{\sigma_T^2}{\sigma_X^2}$ هي ما يسمى بمعامل الثبات ρ فإن

$$\frac{\sigma_E^2}{\sigma_X^2} = 1 - \rho$$

وبالضرب التبادلي ينتج :

$$\sigma_E^2 = \sigma_X^2 (1-\rho) \ldots\ldots\ldots\ldots\ldots\ldots(7)$$

وباستخراج الجذر التربيعي لكل من الطرفين نجد أن :

$$\sigma_E = \sigma_X \sqrt{(1-\rho)} \quad \ldots \quad (8)$$

أي أن الخطـــأ المعياري للقياس =حاصل ضرب الانحراف المعياري للدرجات الملاحظة في الجدر التربيعي لمعامل الثبات مطروحا من الواحد الصحيح .

ويعد الخطأ المعياري للقياس عاملاً أساسيًا في تقرير نتائج الاختبارات والمقاييس وتفسيرها، وهو مرتبط ارتباطاً وثيقاً بمفهوم الثبات ولكنه أيسر فهمــاً لمستخدمي الاختبارات، ولا يتأثر كثيراً بتشتت العلامات على عكس الثـبات. لذلك تهتم أدلة الاختبارات بتحديد القيم التقديرية لهذا الخطأ لكي يتمكــن مستخدم الاختبار من تفسير العلامات تفسيراً احتماليًا. فإذا حصل فــرد في اختبار ما على العلامة 65 فإن هذه العلامة ليست نقطة محددة وإنمــا يمكــن أن تقــع فــي مدى يتراوح بين قيم معينة على متصل العلامــات تحددهــا القيمة التقديرية للخطأ المعياري للقياس ودرجة الثقة المطلوبــة. فإذا كانت قيمة الخطأ المعياري تساوي 1، فبدلا من القول بأن الفــرد حصــل علــى العلامة 65 نقول مثلاً : بأننا نثق 68% أن علامته الحقيقية تقع بين 64، 66 إذا ما أعيد تطبيق الاختبار عليه عدداً كبيراً من المرات(علام،2002).

العلاقة بين الخطأ المعياري للقياس والثبات:

يعتبر الخطأ المعياري للقياس مؤشرا مهما في دقة القياس، ووسيلة من وســائل التنبؤ بقيمة الثبات بحيث أن قيمة الخطأ المعياري تتناسب عكسيا مــع قــيمة معامل الثبات، فإذا كان مقدار هذا الخطأ صفرا(أي القياس تم بدون خطأ) تكون علامات الاختبار متسقة اتساقا تاما أي معامل الثبات = 1، وهذا يعني أنّ العلامة الملاحظة تمثل قدرة حقيقية لجميع المفحوصين.

وبالتالي فإن الارتباط بين العلامات الملاحظة والعلامات الحقيقية = 1، إما إذا كـان معامل الثبات يساوي صفرا، فان العلامة الملاحظة على الاختبار هي علامة عشوائية، والارتباط بين العلامات الملاحظة والعلامات الحقيقية = صفر (عباس،1993) .

وبشكل عام فإن القياس التربوي والنفسي لا يخلو من وجود قدر معين مـن خطـأ القـياس مما يؤدي إلى خفض ثبات درجات الاختبارات بقيم متفاوتة بحسب مقدار هذا الخطأ(علام،2002).

أخطاء القياس:

أخطـاء القياس [illegible] صدق الاختبار وثباته. وتتـنوع مصادرها [illegible] ناتجة لظروف تطبيق الأداة، أو من طبيعة [illegible] (Brown,1983).

وهـناك نوعـان مـن أخطـاء القياس أحدهما: الأخطاء المنتظمة "Systematic Error"، و الآخـر هـو الأخطاء غير المنتظمة أو العشوائية "Unsystematic Error".

أولا: الأخطـاء المنـتظمة "Systematic error" : وهي الأخطاء التي تؤثر بشكل منتظم في عملية القياس أي باتجاه واحد(زيادة أو نقصان) وهي فـي الغالب مرتبطة بأداة القياس وما تتطلبه من إجراءات وعمليات وتؤثر بشكل خاص في صدق الاختبار(ابو زينة، 1992).

ثانياً: الأخطاء غير المنتظمة أو العشوائية "Random error": هي تلك الأخطـاء التـي لا ترتبط بأداء الفرد المستقبلي الذي نود الاستدلال عليه، وبالتالي فهي ترتبط بدرجة الفرد الحقيقية أو بدرجة الأخطاء العشوائية إذا ما أعيد الاختبار نفسه أو صيغة مكافئة له على الفرد نفسه. وتتسم الأخطاء

غـير المنـتظمة(العشوائية) التي تؤثر في درجة الفرد على الاختبار بأنها موقفيه ولا يمكن التنبؤ بكيفية تأثيرها أو مقدارها و لذلك تعمل على خفض ثـبات العلامات(عـلام،2002). ومن هذه الأخطاء العشوائية التخمين والذي يظهر بشكل واضح في اختبارات الاختيار من متعدد .

فالتخميـن هو الأسلوب لذي يلجأ إليه المفحوص للإجابة عن فقرات الاختـبار عـندما لا تكون لديه المعرفة التامة والأكيدة بالإجابة الصحيحة للفقـرة. ففـي مـثل هـذه الحالــة؛ إما أن تكون لدى المفحوص معرفة جزئية(غير تامة أو أكيدة) "Partial Information" تجعله قادرا على حصر الإجابـة الصحيحة بين عدد محدود من البدائل الموضوعة للفقرة، فيختار أحدهـا كإجابـة للفقرة، [illegible] هذه المعرفة الحرئية فيخـتار عشـوائيا من بين جميع البدائل الموضوعة للفقرة أحدها كإجابة للفقرة(سواقد، 1992).

ويمكن تميز نوعين من التخمين هما :-

1. التخميـن الأعمى"Blind Guessing": وهو التخمين العشوائي الذي يتم دون ترجيح لأحد البدائل كأن يمارس الممتحن القرعة عند الإجابة عن الأسـئلة الموضـوعية، وهـنا يكـون احـتمال حصوله على الإجابة الصحيحة هو $\left(\frac{1}{\text{عدد البدائل}}\right)$.

2. التخمين الذكي أو التخمين المبني على معلومات"Informed Guessing" :
وهـو الذي يعتمد على محكات ودلائل تقود إلى الإجابة الصحيحة، قد تكون هذه الدلائل صيغة السؤال أو بعض الكلمات في البدائل المختلفة للسؤال التي تدل على الإجابة الصحيحة، أو يستطيع من خلالها استبعاد بدائل خاطئة وحصر الإجابات الصحيحة بين عدد محدود من الإجابات

ثــم التخميــن بيــن هذا العدد المحدود من البدائل، ولذلك فإن احتمال حصوله على الإجابة الصحيحة يزداد(الخطيب،1993).

وتعــد مشكلة التخمين في الاختبارات الموضوعية من بين المشكلات التــي اهتم خبراء القياس النفسي والتربوي بدراستها وبحثها للتعرف على مــدى تأثيرها سواء في أداء المفحوصين أو على الخصائص السيكومترية لهذه الاختبارات(علام،1990).

فــزيادة التخمين تنقص ثبات الاختبار لأن المفحوص الذي استجاب من خلال التخمين في المرة الأولى لإجراء الاختبار لا يعتمد نفس الأساس في التخمين إذا جلس للاختبار مرة أخرى. وبذلك، تضعف الصلة بين أداء المفحــوص في مر[illegible] قيمة معامل الثبات. وتحــتلف الاختــبا[illegible] تبعاً لنوعها. وتعتبر اختبارات الاختيار م[illegible] تأثراً بالتخمين تليها اختــارات الاختــيار من متعدد بحيث يقل تأثرها بالتخمين مع زيادة عدد البدائل للفقرة(ملحم، 2000).

وللتقلــيل مــن هذه المشكلة، قدم العاملون في ميدان القياس والتقويم الــتربوي طرقاً ومعالجات إحصائية، اعتمدت جميعها على طريقة تصحيح الاختــبار والتعلــيمات المنسجمة معها. تلك المعالجات تحدد القاعدة التي تعطــى على أساسها العلامة التي تحدد للمفحوص بحيث تعبر عن مستوى أدائه بدقة اكبر(سوالمة، 1992).

فيمكــن أن تصــحح أســئلة الاختبارات بطرق مختلفة. وعلى جميع الأحــوال فإن العلامة التي يحصل عليها المفحوص تمثل نوعاً من التقدير لمستوى قدرته في السمة المقاسه، والشائع في تصحيح فقرات الاختبار من متعدد هو إعطاء قيمتين للاستجابة باعتبارها متغيراً ثنائياً(Dichotomous)؛

أي تعطـــى الإجابة عن السؤال(1) إذا كانت صحيحة، وتعطى العلامة (0) إذا كانت خطأ.

إن هــذه الطــريقة لتصــحيح الاحتــبارات، تقســم المفحوصين إلى مجموعتيــن: مجموعـــة الإجابـــات الصحيحة وهؤلاء إما اختاروا الإجابة الصـــحيحة عـــن معـــرفة تامة أو حصلوا على الإجابة الصحيحة بواسطة التخميـــن، سواء كان لديهم معرفة جزئية أو لم يكن، ومجموعة الإجابات الخاطـــئة وهـــؤلاء إمـــا يملكون معلومات غير صحيحة تماما أو لديهم معلومـــات غـــير صحيحة جزئياً فيختارون أحد البدائل الخاطئة باستخدام التخمين وقد يكون بعضهم قد اخطأ في اختيار البديل الصحيح لسبب ما مع انهم يعرفونه وهكذا فان الفقرة تكون خاطئة مع انها لم تكن نتيجة التخمين.

وعلـــيه فإنهـــا غير قادرة على تدريج المفحوصين وفق ما لديهم من معـــرفة،على متصـــل يمـــتد مـــن المعـــرفة الكاملة إلى المعلومات غير الصـــحيحة تمامـــا، كما انها غير قدرة على ضبط أثر التخمين مما دفع بالمهتميـــن بالبحـــث عـــن طـــرق أخـــرى تأخذ بعين الاعتبار مستويات المعـــرفة عـــند المفحوصـــين، وتكون قادرة على قياس من أثر التخمين (Jaradat & Tollefson,1988). ومن طرق التصحيح المستخدمة في تصحيح اختبارات الاختيار من متعدد من أثر التخمين الطرق الآتية:

الطريقة الأولى : معادلة التصحيح لأثر التخمين(العقاب)

يعـــتمد نموذج الاستجابات العشوائية model for random responses علـــى افتراض ان المفحوص في الاختبارات الموضوعية إما أن يستجيب اســـتجابة صحيحة على الفقرة مع تأكد تام؛ أي عن معرفة حقيقية، أو انه يقـــوم بتخميـــن عشوائي ويكون احتمال حصوله على الاستجابة الصحيحة

يساوي $\left(\frac{1}{\text{عدد البدائل}}\right)$. ففي الاختبارات من نوع(صح،خطأ) سيكون احتمال الحصول على إجابة صحيحة بالصدفة 0.5، إما في اختبارات الاختيار من مستعدد ذات الأربـــع بدائل فإن الاحتمال سيكون 0.25. وهكذا فإن علامة المفحــوص ســتتكون من عدد الإجابات الصحيحة التي حصل عليها عن معــرفة تامة بالإضافة إلى عدد الإجابات الصحيحة التي حصل عليها عن طــريق التخمين، لكن من الصعب قياس عدد الاستجابات التي تمثل معرفة تامة ولكن نستطيع أن نقدرها فإذا رمزنا لـــ.

عدد الاستجابات الصحيحة برمز R

عدد الاستجابات الخاطئة برمز W

عدد الفقرات التي حاولها المفحوص برمز T حيث $T = R + W$

عدد البدائل للفقرة برمز K

عدد الاستجابات الصحيحة التي تمثل معرفة تامة برمز R_C

فتكون عدد الاستجابات الصحيحة يساوي

$$R = R_C + P(T - R_C)\text{........................}(9)$$

حيث P احتمالية حصول الفرد على إجابة صحيحة = $\frac{1}{K}$

و تعنـــي هذه المعادلة أن علامة المفحوص R ناتجة من جزء ين R_C وهـــي العلامة الحقيقية و $P(T - R_C)$ الناتجة عن التخمين، فإذا حاول الفرد الإجابـــة عـــن كل الفقرات في الاختبار، فان الحد الأخير في المعادلة(9) يصبح $(N - R_C)$ حيث N عدد فقرات الاختبار. وبتبسيط المعادلة 9 ينتج

$$R = R_C + PT - PR_C$$

وبترتيب الحدود و إخراج عامل مشترك ينتج :

$$R = R_C - PR_C + PT$$

$$R = R_C(1 - P) + PT$$

$$R_c(1-P) = R - PT$$
$$= R - P(R+W) = R - PR - PW$$
$$= R(1-P) - PW$$

وبقسمة طرفي المعادلة على $(1-P)$ نجد:

$$R_c = \frac{R(1-P)}{(1-P)} - \frac{PW}{(1-P)}$$

وبتعويض بدل P قيمتها $\frac{1}{K}$ ينتج :

$$R_c = R - \frac{W}{1-K} \qquad \text{..........................(10)}$$

وتعـــرف هذه المعادلة بمعادلة التصحيح لأثر التخمين(العقاب) حيث R_C ترمز للدرجة الكلية في الاختبار أو العلامة الكلية بعد التصحيح من أثر التخميــن، R ترمـز [illegible]، W ترمز إلى عدد الإجادت الخاطئة(8[illegible])

وتفترض هذه [illegible] الاحتمال الأول أن المفحوص يعرف الإجابة الصحيحة للفقرة وعليه فإن احتمال حصوله عليها يســاوي 1، والاحتمال الثاني إن المفحوص لا يعرف الإجابة الصحيحة للفقــرة وبالتالــي فإنه يجيب بطريقة عشوائية عن هذه الفقرة .وهكذا فان احــتمال حصوله على الإجابة الصحيحة = $\frac{1}{\text{عدد البدائل}}$ ، كما أن احتمال حصـــوله علـــى الإجابـــة غـــير الصـــحيحة = $\frac{\text{عدد بدائل الفقرة} - 1}{\text{عدد البدائل}}$

(Ineke & Wilfried & Paul,2002).

وبـــناءاً علـــى ذلك فان استخدام هذه الصيغة يقتضي خصم جزء من العلامة الكلية للمفحوص بناطر ما يفترض أنه اعتمد فيه على العشوائية في الإجابـــة، ولا تدخـــل في الاعتبار المعلومات الجزئية التي ربما يكون قد استند إليها في إجابته(علام،1990).

وبمـــا أن مقـــدار ذلك الجزء غير معروف، فانه يتم تقديره من عدد الإجابات غير الصحيحة من خلال الاحتمال النظري، فيكون مقابل عدد من الإجابـــات غير الصحيحة مساوٍ لـــ(عدد بدائل الفقرة – 1) هناك إجابة واحـــدة صـــحيحة، حصـــل عليها المفحوص باستخدام التخمين العشوائي، وعليه فان علامة المفحوص على الاختبار تصبح مساوية

لـــ : عدد الإجابات الصحيحة – (عدد الإجابات غير الصحيحة التي حاولها) / (عدد بدائل الفقرة – 1) .

"ويـــرى لورد ونوفك (Lord & Novick) انه بدون هذه الافتراضات يصـــبح مـــن الصـــعب إيجـــاد وســـيلة لضـــبط التخمين، و يشير ننلي (Nunnally,1978) إلى ان استخدام المعادلة يميل لمعاقبة المفحوصين ذوي المعـــرفة الجزئـــية اكـــثر ممـــا يجـــب، في حين يميل استخدامها لمعاقبة المفحوصـــين ذوي المعلومات غير الصحيحة اقل مما يجب. غير ان لورد Lord يقول :" إذا كانت طريقة التصحيح تتضمن معاقبة بدل تخمين، وأراد المفحوص أن يرفع علامته إلى الحد الأقصى، فمن [illegible] يجيب عن كل فقرة يشعر ان لديه عنها معرفة جزئية ترشده لاختيار [illegible] صحيح، أو تمكنه من استبعاد أياً من البدائل التي يعتقد حازماً [illegible] غير صحيحة، ثم يقوم بالتخمين من بين ما تبقى من البدائل. وبناءاً على [illegible] تكون تعليمات الاختبار واضحة بما فيه الكفاية لإعلام المفحوص [illegible] باتباع هذه الاستراتيجية عند اخذ الاختبار. مثل هذه التعليمات [illegible] تعليمات معادلـــة التصـــحيح لأثـــر التخمين ". ويشير لورد إلى ان استخدام معادلة التصـــحيح وفـــق هـــذه التعلـــيمات يمكن المفحوص من استغلال المعرفة الجزئـــية لديـــه، إذ انـــه سيحذف فقط الفقرات التي يعتقد أن إجاباته عنها ستكون بمجرد التخمين العشوائي"(سواقد،1992).

وبناءاً على ذلك فقد توصل لورد إلى الافتراض الأتي :" إن الفرق بين ورقة اختبار صححت وفق تعليمات معادلة التصحيح لأثر التخمين، ونفس الورقة إذا صححت تحت تعليمات إجابة جميع الفقرات، هو فقط في أن الفقرات المحذوفة إن وجدت في الحالة الأولى، قد استبدلت بالتخمين العشوائي في الأخيرة. وهذا يعني أن المفحوص إما يعرف الإجابة أو يخمن بشكل عشوائي (Bliss,1980).

وقد تم احتبار هذا الافتراض من قبل العديد من الدراسات ومنها دراسة بلس(Bliss,1980) التي أشارت إلى عدم صحة هذا الافتراض، وتوصل من خلالها إلى أن المفحوصين الذين يأخذون الاختبار تحت تعليمات معادلة التصحيح لأثر التخمين (العقاب)، لا يحذفون الفقرات التي تكون إجاباتهم عليها هي مجرد تخمين عشوائي فقط، بل يميلون إلى حذف الفقرات التي قد يكون احتمال إجاباتهم عنها صحيحة أفضل من الاحتمال النظري للتخمين العشوائي.

وبالرغم من أن الافتراضات السابقة الخاصة بمعادلة التصحيح لأثر التخمين القائمة على العقاب ضرورية للحصول على وسيلة لتقدير أثر التخمين، إلا أن دراسات عديدة تناولت أثر استخدام هذه المعادلة، قد شككت بمدى نجاعتها في تحسين الخصائص السيكومترية للاختبار بشكل عام. إذ يشير البعض إلى أن استخدام هذه المعادلة يعاقب المفحوصين الأقل ميل للمخاطرة أكثر من زملائهم الأكثر ميلا للمخاطرة كما يعاقب المفحوصين ذوي التحصيل العالي أكثر مما يستحقون، ويعزو لورد النتائج التي توصلت إليها تلك الدراسات، إلى قصور تعليمات الاختبار في إقناع المفحوصين باتباع الاستراتيجيات الصحيحة عند أخذ الاختبار، والتي

تجعلهم يحذفون فقط الفقرات التي تكون إجابتهم عليها اكثر من مجرد التخمين العشوائي(Jaradat & Sawaged ,1986)

ولا تزال هذه الصيغة مثار حوار ودراسة من جانب خبراء القياس، فالمؤيدون منهم يرون أن تصحيح الدرجات الكلية من أثر التخمين باستخدام هذه الصيغة تجعل العلامات الناتجة أكثر اقترابا من العلامات الحقيقية للمفحوصين، وتحثهم على الابتعاد عن التخمين مما يؤدي إلى تحسين ثبات الاختبارات الموضوعية إما المعارضون فيرون أن ترتيب المفحوصين بالنسبة لبعضهم البعض اعتمادا على درجات الاختبار لا يتغير سواء استخدمت الدرجات الكلية دون تصحيحها من أثر التخمين أو بعد أجراء هذا التصحيح [illegible] أثر التخمين المبني علــى معلومات جزئية [illegible] افتراض خاطئ مؤداه أن العلامات بعد تصحيحها [illegible] المعلومات الحقيقية(عــلام 1990).

ويقترح أصحاب هذا التوجه بعض الإرشادات التي يمكن أن تسهم في تقليل أثر التخمين بقدرٍ كبير كالتالي :

1- تحذير المفحوصين من التخمين العشوائي في إجابتهم عن فقرات الاختبار.
2- السماح بوقت كاف لكل فرد لكي يتمكن من محاولة إجابة كل فقرة من فقرات الاختبار.
3- تضمين البدائل التي تشتمل عليها فقرات الاختيار من متعدد معلومات صحيحة جزئياً أو أخطاء شائعة تتعلق بالمجال الذي تقيسه الفقرات.

ويؤكـد لورد انه إذا كانت تعليمات الاختبار تشجع المفحوصين على عـدم تـرك أي فقـرة دون إجابـة، فان العلاقة بين العلامات الكلية قبل التصـحيح مـن اثر التخمين وبعد تصحيحها تكون علاقة تامة؛ أي معامل الارتباط بينهما يساوي الواحد الصحيح(علام، 1990).

إما فيما يتعلق بتأثير استخدام معادلة التصحيح لأثر التخمين(العقاب) علـى ثـبات الاختـبار، فإنـنا نجد كثيراً من التعارض في وجهات نظر المهتمين بالموضوع. ففي حين يشير البعض إلى أن معامل ثبات الاختبار يقـل باسـتخدام معادلـة التصحيح مقارنةً بمعامل الثبات المستخرج عند التصـحيح بالطريقة التقليدية القائمة على أساس عدد الإجابات الصحيحة. يؤكـد البعض الآخر أن استخدام هذه المعادلة يزيد معامل الثبات للاختبار (سواقد،1983).

الطريقة الثانية : معادلة التصحيح لأثر التخمين(المكافأة)

جـاءت هذه الطريقة كبديل للطريقة السابقة وتم اقتراحها بواسطة (Traub & Hambleton & Singh) عـام (1969)، وحققـت نفس هدف الطـريقة السـابقة عـن طريق تعويض المفحوص عن الفقرات المحذوفة بمقدار $\frac{1}{K}$ من العلامات عن كل فقرة محذوفة، وعدم اتخاذ أي إجراء فيما يتعلق بالإجابات الخاطئة(Budescu & Bar-hillel ,1993).

وتعـتمد هـذه الطريقة على النموذج الذي يأخذ بعين الاعتبار ثلاث مواقف ممكنة لاستجابات الأفراد :

1- أن المفحوص يعرف الإجابة الصحيحة واختارها.

2- أن المفحوص حذف الفقرة ولم يجب عليها.

3- أن المفحــوص خمن بشكل عشوائي(واختار أحد الفقرات أجابها بشكل عشوائي).

وبالاعــتماد على هذا النموذج(نموذج التخمين العشوائي) فان الصيغة الأساسية لتصحيح العلامات من اثر التخمين تعطى بالمعادلة التالية :

$$R_c = R + \frac{O}{K} \ldots \qquad (11)$$

حيث R_c ترمز للدرجة الكلية في الاختبار

R ترمز إلى عدد الإجابات الصحيحة.

O ترمز إلى عدد الفقرات المحذوفة(أي التي تركها بدون إجابة)

K ترمز إلى عدد البدائل التي تشتمل عليها الفقرة(Crocker & Algina,1986).

وعرفــت هــذه الطريقة باسم طريقة معادلة تروب(Traub) وتقوم على تشجيع المفحوص على عدم التخمين وذلك بمكافئته إذا لم يخمن، وذلك بــزيادة علامــته الظاهرية بما يتناسب مع درجة تحفظه على الإجابة، أو بابتعاده عن الإجابة عندما لا يرجح أي بديل منها. وذلك من خلال أضافه ذلــك الجــزء مــن العلامة الذي يتوقع أن يحصل عليه فيما لو خمن إلى علامته الظاهرية حسب الصيغة الرياضية السابقة.

ونلاحــظ أن هــذه الطريقة من التصحيح تزيد من علامة المفحوص الظاهــرية بواســطة مكافأتــه بنقاط إضافية على الفقرات المحذوفة على افتراض انه إذا كان المفحوص قد حاول الإجابة على الفقرة المحذوفة فان احتمالــية اختــياره للإجابة الصحيحة يساوي $\left(\frac{1}{\text{عدد البدائل}}\right)$، وتعتمد على افــتراض أن جميع محاولات التخمين على الفقرات المحذوفة سوف تكون عشوائية(Crocker & Algina, 1986).

ويرى تروب أن هذه المعادلة تقوم على أساس نفسي يختلف عن المعادلة السابقة فهي تعد بإضافة درجات عن الأسئلة المتروكة بدلاً من خصم درجات عن الإجابات الخاطئة، وإذا ذُكر هذا بوضوح في تعليمات الاختبار، فإنها تكون أقل إرهاباً للتلميذ الذي يستسلم بسهولة للتهديد بالعقاب (Nunnally 1978; Crocker & Algina, 1986)

ان الفلسفة التي تقوم عليها هذه الطريقة ترى ان المفحوص إذا خمن سيجيب نصف الأسئلة إجابة صحيحة إذا كانت من نمط صحيح – خاطئ، وعلى ثلث الأسئلة إذا كان عدد بدائل إجابة السؤال ثلاثة وعلى ربع الأسئلة إذا كان عدد البدائل أربعة. ولكي نشجعه على عدم التخمين نرى ان نضيف لعلامته [illegible] بواسطة التخمين. وهذا ما يقودنا إلى تطبيق المعادلة رقم(11) بدلا من المعادلة رقم(10). وفي هذه الحالة نحصل على علامات متوسطها وتباينها أعلى من متوسط العلامات الناتجة من تطبيق المعادلة رقم(10)، ولكن ترتيب المفحوصين سيكون واحد في الحالتين(أبو لبدة، 1985).

يقول ايبل (Ebel,1972) انه إذا جرى تصحيح مجموعة من العلامات لأثر التخمين بطريقتين مختلفتين، بطرح كسر (Fraction) من الإجابات الخاطئة(أي استخدام معادلة معاقبة التخمين)، وأضافه كسر (Fraction) من الإجابات المحذوفة(Omitted Answers)؛ أي استخدام معادلة المكافأة. يمكن الحصول على مجموعتين مختلفتين من العلامات المصححة، وعلى الرغم من أن مجموعتي العلامات سوف تختلفان في قيمتي متوسطيهما (العلامات المصححة على أساس الإجابات المحذوفة تكون أكثر تباينا في كل الحالات)، فإن مجموعتي العلامات ترتبطان ارتباطا تام.

وفــيما يتعلق بتأثير استخدام هذه المعادلة على ثبات الاختبار، فإن بعض الدراسات النظرية تشير إلى أن استخدام المعادلة التي تتضمن مكافأة عــدم التخميــن تعطـــي معـــامل ثـــبات عالٍ إذا ما قورن بمعامل الثبات المستخرج عند توظيف معادلة عقاب التخمين(Budescu & Bar-hille, 1993).

الطريقة الثالثة : معادلة التصحيح لأثر التخمين(المكافأة والعقاب)

تصـــحيح أثر التخمين (Correction for guessing) يستند إلى الافتراضات التالية :

1. أن المفحوص يعرف الإجابة الصحيحة
2. أن المفحوص لا يعرف الإجابة الصحيحة، ويحذف الفقرة.
3. أن المفحـــوص لا يعرف الإجابة على الفقرة وبالتالي يجيب على الفقرة بشكل عشوائي (Ben[illegible]).

الطرق السابقة لم تكن تعتبر جميع الافتراضات السابقة. فنجد ان الطـــريقة الأولى اعتمدت على الإجابات الصحيحة والخاطئة، والطريقة الثانية اعتمدت على الإجابات الصحيحة والإجابات المحذوفة. فجاءت هذه الطـــريقة لتجمع الافتراضات الثلاثة السبقة بمعادلة واحدة تعرف بمعادلة التصحيح لأثر التخمين(المكافأة والعقاب).

وتمتاز هذه المعدلة عن سابقتيها بأنها تكافئ المفحوص على الفقرات المتروكة بدون إجابة مبنية على التخمين، وتعاقبه على الفقرات التي يُتوقع أنـــه قـــد خمـــنها، أي تجمع بين المكافأة والعقاب وتنص على ان العلامة المصححة بهذه الطريقة تساوي

$$\text{عدد الإجابات الصحيحة} + \frac{\text{عدد الفقرات المحذوفة}}{\text{عدد البدائل}} - \frac{\text{عدد الإجابات الخاطئة}}{\text{عدد البدائل} - 1}$$

ومـيزة هـذه المعادلـة أنها توازن بين المبالغة في درجة المفحوص الناتجة عن المكافأة، والإبخاس الناتج عن العقاب.

الطريقة الرابعة : الحذف الجزئي "Subset elimination"

تـم اقـتراح هـده الطريقة من قبل كومبس(Coombs) سنة 1953، وتنص تعليمات هذه الطريقة ان يستبعد المفحوص البدائل غير الصحيحة، حيـث يكافـئ المفحـوص بعلامة واحدة لكل بديل غير صحيح يستبعده، ويعاقب بخصم(n-1) من العلامات، حيث n تمثل عدد البدائل للفقرة، لكل بديل صحيح يستبعده(Bush,2001).

وهذه الطريقة تميز بين المستويات المختلفة للمعرفة عند المفحوصين فالمفحوص الذي يمتلك معرفة تامة "Full knowledge" يستبعد جميع البدائل الخاطئة، والمفحوص الذي يمتلك معرفة جزئية "Partial knowledge"يستبعد مجموعـة جزئية من البدائل الخاطئة، والمفحوص الذي يمتلك معلومات غير صحيحة جزئيا "Partial misinformation" يستبعد الإجابة الصحيحة ومجموعة جزئية من البدائل الخاطئة،والمفحوص الذي لا يمتلك المعلومات تمام"Full misinformation" يستبعد الإجابة الصحيحة فقط (Ban-simon & Budescu,1997).

الطريقة الخامسة : الاختيار الجزئي للبدائل "Subset selection"

وتكـون آلـية التصـحيح بهذه الطريقة ان يعطى المفحوص علامة مقدارها (عدد بدائل الفقرة – 1)، عن كل بديل صحيح تتضمنه المجموعة الجزئية التي يختارها المفحوص، وتخصم منه علامة واحدة عن كل بديل غير صحيح تتضمنه المجموعة الجزئية التي يختارها المفحوص.والجدول

رقم(1) يوضح طريقة التصحيح لفقرة اختيار من متعدد ذات خمس بدائل بهذه الطريقة (Gibbons & Olkin & Sobel 1979 ,Frary 1988).

جدول(1)

طريقة تصحيح فقرة اختيار من متعدد ذات خمس بدائل باستخدام طريقة الاختيار الجزئي

المجموعة الجزئية	علامة المجموعة التي تحتوي على الإجابة الصحيحة	علامة المجموعة التي لا تحتوي على الإجابة الصحيحة
0		0
1	4	1-
2	3	2
3	2	3-
4	1	4-
5		-

الدراسات السابقة:

استحوذ التخمين والطرق المستخدمة لضبطه على اهتمام العديد من الباحثين، وقد حاولت هذه الدراسات؛ استقصاء فعالية هذه الطرق في ضبط التخمين من جهة، وكيف يؤثر ذلك كله على الخطأ المعياري للقياس ومعامل ثبات الاختبار وخصائص توزيع العلامات.

الدراسات المتعلقة باستخدام معادلة التصحيح لأثر التخمين(العقاب).

"يلخص راش Rasch نتائج الدراسات المتيسرة لديه والتي استخدمت فيها معادلة التصحيح لأثر التخمين ويستنتج أن الدلائل تشير إلى إمكانية الحصول على علامات اكثر صدقا، عندما يطلب من المفحوصين حذف الفقرات التي لا تكون إجاباتهم عليها اكثر من مجرد التخمين العشوائي أما فيما يتعلق بدلالات ثبات الاختبار، فالأمر متنازع عليه بين من يقول انها تزيد، ومن يقول انها تنقص من معامل ثبات الاختبار"(سواقد،1983).

وفي ضوء ذلك أجريت العديد من الدّراسات لاختبار اثر هذه الطريقة علـــى معـــامل ثـــبات الاختـــبار ومـــن هـــذه الدّراسات. دراسة النبهان (Alnabhan.2002) تـــناول فـــيها اختبار ثلاث طرق لتصحيح أثر التخمين على الخصائص السيكومترية للاختبار حيث تمت المقارنة بين طريقة عدد الإجابـــات الصـــحيحة وطـــريقة اســـتخدام معادلـــة التصـــحيح لأثـــر التخميـــن(العقـــاب) وطريقة الاختيار الجزئي، ولتحقيق هدف الدراسة تم اعـــداد اختـــبار تحصيلي من نوع الاختيار من متعدد لقياس المصطلحات والمفاهـــيم في مادة مناهج البحث في علم النفس، وتم تطبيق الاختبار على أفـــراد العيـــنة الـــبالغ عددهم(120) طالبا من الطلبة المسجلين في مبحث "مـــناهج البحـــث في علم النفس [illegible] في من العام الدراسي 2000\2001، أظهـــرت [illegible] لتأثير معادلة التصحيح لأثر التخمين(العقاب) وطريقة الاختيار الجزئي على معامل ثبات الاختبار تأثـــير متســـاوي، وتعطـــيان معامل ثبات أعلى من الطريقة الأولى وهي طريقة عدد الإجابات الصحيحة.

وفـــي دراســـة لـــ(Muijtjens etal ,1999)أجرى فيها مقارنة بين طريقة عـــدد الإجابـــات الصـــحيحة(Number – Right) وطـــريقة استخدام معادلة التصـــحيح لأثر التخمين(العقاب) وبينت نتائج هذه الدراسة إلى أن معامل ثبات الاختبار يكون أعلى عند تصحيح الاختبار باستخدام معادلة التصحيح لأثر التخمين(العقاب) مقارنة بطريقة عدد الإجابات الصحيحة.

كما أجرى جرادات وتولفن(Jaradat & Tollefson ,1988) دراسة هدفت إلـــى الـــتّعرف على أثر طرق تصحيح اختبارات الاختيار من متعدد على الخصـــائص الســـيكومترية للاختـــبار. حيث تم المقارنة بين ثلاث طرق للتصـــحيح هـــي معادلة التصحيح لأثر التخمين(العقاب) وطريقة الاختيار

الجزئـــي مـــن بيـــن محموعة البدائل الموضوعة للسؤال وطريقة استبعاد الـــبدائل غـــير الصـــحيحة. وتكونت عينة الدراسة من (54) طالبا والّذين يدرســون مادة القياس النفسي، وأظهرت النتائج فيما يخص معامل الثبات تفوق طريقة الاختيار الجزئي.

وفي دراسة أخرى لـــ جرادات و سواقد (Jaradat & Sawaged ,1986) تناولـــت اختـــبار ثلاث طرق لتصحيح اختبارات الاختيار من متعدد لأثر التخمين. وأثرها، على الخصائص السيكومترية للاختبار عند المفحوصين من مستويات تحصيل، ودرجات مخاطرة مختلفة. والطرق الثلاث لتصحيح الاختبار هي:

1- الطـــريقة التقلـــيدية 2- معادلـــة التصـــحيح لأثر التخمين(العقاب)

3 طريقة الاختيار الجزئي من بين محموعة البدائل الموصوعة للسؤال.

وقـــد اعـــد ثلاثـــة أشكال متكافئة لاختبار، يقيس تحصيل التلاميذ في مبحـــث العلـــوم العامـــة للصف الثالث إعدادي، أخضعت لها عينة مكونة من (173) طالبا وطالبة، صحح كل شكل منها بإحدى الطرق الثلاث، وأظهرت النـــتائج فـــيما يـــتعلق بأثر طريقة تصحيح الاختبار على معامل ثباته، ان تصحيح الاختبار باستخدام معادلة التصحيح لأثر التخمين (العقاب) و بطريقة الاختيار الجزئي يعطي معامل ثبات أعلى مقارنة بتصحيح الاختبار بطريقة التقليدية(تصحيح الاختبار على أساس عدد الإجابات الصحيحة).

كما وأجرى بلس(Bliss,1980) دراسة هدفت إلى فحص فرض لورد المتعلق بسلوك التخمين عند المفحوصين والذي ينص على- إن الفرق بين ورقة اختبار صححت وفق تعليمات معادلة التصحيح لأثر التخمين، ونفس الورقـــة إذا صـــححت تحت تعليمات إجابة جميع الفقرات، هو فقط في أن الفقـــرات المحذوفـــة إن وجـــدت في الحالة الأولى، قد استبدلت بالتخمين

العشـوائي فـي الأخيرة. وهذا يعني أن المفحوص إما يعرف الإجابة أو يخمـن بشـكل عشـوائي - على اختبارات الاختيار من متعدد عند طلبة المـرحلة الابتدائية،حيث استخدم اختباراً تحصيلياً لمادة الرياضيات مكون مـن(26) فقـرة موضوعية من الاختيار من متعدد، وكانت عينة الدراسة (168) طالـبا مـن طلاب المرحلة الابتدائية، فقد أعطي المشتركون في الدراسـة الاختـبار تحت تعليمات معادلة التصحيح لأثر التخمين وبعد ان أنهى جميع المفحوصين الإجابة عن الاختبار، أعيدت إليهم أوراق الإجابة الأولـى مـع أقـلام ملونة، وطلب إليهم محاولة الإجابة عن الفقرات التي حذفوها في السابق تحت تعليمات عدم التخمين مع التوضيح بان علاماتهم ستحسب على أساس عدد الإجابات الصحيحة. وأشارت نتائج هذه الدراسة إلـى ان معـامل ثـبات الاختبار يكون أفضل في حال تصحيح الاختبار باستخدام معادلة التصحيح لأثر التخمين مقارنة بطريقة التصحيح التقليدية.

وفي دراسة كروس و فراري(Cross & Frary ,1977) هدفت إلى اختبار نـتائج لـورد المـتعلقة بتصحيح اختبارات الاختيار من متعدد، وتم فيها المقارنـة بيـن طريقتين لتصحيح الاختبار طريقة تصحيح الاختبار على أسـاس عـدد الإجابات الصحيحة، وطريقة استخدام معادلة التصحيح لأثر التخميـن(العقـاب). أظهرت النتائج ان تصحيح الاختبار باستخدام معادلة التصحيح لأثر التخمين يعطي معامل ثبات أعلى.

إما بخصوص طريقة استخدام معادلة التصحيح لأثر التخمين (المكافأة) وطريقة معادلة التصحيح لأثر التخمين(المكافأة والعقاب)، فلم تتوفر - في حـدود معـرفة الباحث - دراسات عملية لاختبارها، ومقارنة نتائجها مع غـيرها من الطرق الممكنة لتصحيح اختبارات الاختيار من متعدد وهو ما تحاول هذه الدراسة القيام به.

و إمــا فـيما يـتعلق بأثر طرق التصحيح على شكل التوزيع لعلامات الاختــار فلــم تتوفر دراسات حول هذا الموضوع وبالتالي ستحاول هذه الدراسة القيام بذلك.

من خلال الاستعراض المختصر لمجموعة من الدراسات التي تناولت ضـــبط أثر التخمين في الاختبارات نلاحظ أن غالبية الدراسات السابقة قد أجمعــت علـــى اســتخدام معادلة التصحيح لأثر التخمين(العقاب) وطريقة الاختيار الجزئي مقابل طريقة التصحيح التقليدية، فهناك دراسات أظهرت وجــود فــروق ذات دلالـــة معــنوية بين معاملات الثبات تعزى لطريقة التصــحيح ولصـــالح طــريقة التصحيح باستخدام العقاب مقارنة بالطريقة التقليدية، مثل دراسة بلس(Buss.1950) ودراسة كروس و فراري(& Cross Frary, 1977). وهناك بعض الدراسات التي أظهرت عدم وجود فروق ذات دلالـــة إحصـــائية في معاملات الثبات تعزى لطريقة التصحيح باستخدام العقاب، أو لطريقة الاختيار الجزئي، مثل دراسة جرادات و سواقد (Jaradat & Sawaged, 1986) ودراسة النبهان (Alnabhan.2002).

ونظــراً لعــدم وجــود الدراســات تتعلق بمقارنة هذه الطرق بطرق تصـــحيح أخـــرى مثل طريقة المكافأة وطريقة المكافأة والعقاب فإن إجراء المزيد من الدراسات والتي تأخذ بعين الاعتبار الطرق المختلفة لضبط أثر التخمين، وعلى عينات مختلفة، يعد غاية من الأهمية سعياً وراء الوصول إلى طريقة التصحيح التي تجعل علامة المفحوص على الاختبار اكثر دقة فـــي تمثـيل قدرته الحقيقية على السمة موضوع القياس، وذلك في ضوء قدرتهـــا على ضبط التخمين. وعليه فإن هذه الدراسة تعد مساهمة في هذا المجال.

الفصل الثالث

المنهجية والإجراءات

يتناول هذا الفصل وصفاً للطرق والإجراءات التي اتبعت في تنفيذ هذه الدراسة من حيث عينة الدراسة، وأدوات الدراسة(صدقها وثباتها)، والمعالجات الإحصائية التي اتبعت لتحليل البيانات.

عينة الدراسة:

تكونت عينة الدراسة من جميع الطلاب الذين درسوا مادة مبادئ في القياس والتقويم بكلية التربية في جامعة مؤتة في الفصل الدراسي الثاني 2003\2002، وقد بلغ عددهم([illegible]) طالباً وطالبة موزعين على خمس شعب، وقد تم اختيار شعبة بطريقة عشوائية من بين الشعب الخمسة وكان عدد المسجلين فيها ([illegible]) طالباً وطالبة شكلوا عينة الدراسة الاستطلاعية . والباقي وعددهم(221) شكلوا عينة الدراسة التجريبية.

أداة الدّراسة:

تم تطوير اختبار تحصيلي يشتمل على(60) فقرة من نوع الاختيار من متعدد ذات أربع بدائل لقياس المصطلحات والمفاهيم الأساسية في مقرر " مبادئ القياس والتقويم " الذي يُدرَّس لطلاب كلية التربية بجامعة مؤتة. وذلك بالاستفادة من الاختبار الذي طوره(علام،2001) كما ظهر في الملحق رقم (1).

بعد ذلك تم عرض فقرات هذا الاختبار على سبعة من أساتذة علم النفس(خمسة منهم يُدّرسون مادة مبادئ في القياس) لمراجعة فقرات الاختبار للتحقق من وضوح عباراتها وسلامة صياغتها واستيفائها لشروط

هـــذا النوع من الفقرات وملائمة الفقرات للمحتوى المقرر. وبعد مراجعة الفقرات بناءاً على مقترحات المحكمين ومناقشتهم تم تعديل بعض فقرات الاختــبار وحذف مجموعة أخرى لعدم مطابقتها لمحتوى الخطة الدراسية لهــذا المقـــرر فـــي جامعة مؤتة(ملحق رقم3)، ليتكون الاختبار في شكله النهائي من (30) فقرة(ملحق رقم 2).

الإجراءات:

أولا: تطبيق الاختبار التحصيلي على عينة الدراسة الاستطلاعية :

طُــبِّق الاختــبار علـــى أفـــراد عينة الدراسة الاستطلاعية في ضوء تعليمات تنص على أن تصحيح الاختبار سوف يكون على أساس مجموع الإجابــات الصـــحيحة، وحســـب الوسط الحسابي، والانحراف المعياري، ومعـــامل الثبات باستخدام معادلة كرونباخ(α) والجدول رقم(2) يبين نتائج المفحوصين في عينة الدراسة الاستطلاعية على الاختبار التحصيلي.

الجدول رقم(2)

الوسط الحسابي والانحراف المعياري ومعامل الثبات للاختبار التحصيلي، عند عينة الدراسة الاستطلاعية

عدد المفحوصين	الوسط الحسابي	الانحراف المعياري	معامل الثبات
55	15 5	5 68	0 82

ويتضـــح من الجدول أن معامل ثبات الاختبار هو(0.82) وهو مؤشر كافـــي لاســـتخدام هذا الاختبار لأغراض الدراسة، خاصة إذا أخذنا بعين الاعتبار كون الاختبار قصير وعدد فقراته(30).

ثانياً:تطبيق الاختبار على عينة الدراسة التجريبية

1- تــم تقســيم فقــرات الاختبار إلى ثلاث مجموعات، بحيث تشتمل كل مجموعــة منها على(10) فقرات. رتبت أسئلة هذه المجموعات الثلاث في ثلاثة نماذج.

2 تم توزيع النماذج الثلاثة من الاختبار عشوائيا حسب الطريقة التي سيتم بها تصحيح كل نموذج فكانت على النحو التالي:

النموذج الأول : استخدام معادلة التصحيح لأثر التخمين(العقاب) والتعليمات المنسجمة معها.

النموذج الثاني: استخدام معادلة التصحيح لأثر التخمين(المكافأة) والتعليمات المنسجمة معها.

النموذج الثالث : استخدام معادلة التصحيح لأثر التخمين(المكافأة والعقاب) والتعليمات المنسجمة معها.

3- قُسِّم افراد الدراسة عشوائيا إلى ثلاث مجموعات.

4- أعطيــت كــل مجموعــة واحــدا من نماذج الاختبار مع التعليمات المنســجمة مــع طــريقة تصــحيحه، وتــم توضيح التعليمات للمشــتركين، كما تم تذكيرهم بالاستراتيجية التي يفضل اتباعها عــند اخذ كل نموذج،وبعد التأكد من فهم المشتركين للتعليمات، طلب إليهم البدء بالإجابة على الاختبار.

5- وبعــد أســبوع، تم تطبيق نموذج آخر من الاختبار على كل مجموعة بحيث يكون مختلفا عن النموذج الذي تم تطبيقه قبل أسبوع.

6- بعد مرور أسبوع آخر. تم تطبيق النموذج الأخير لكل مجموعة.

وبذلك تكون المجموعات الثلاث قد جلست للاختبار ثلاث مرات لاستكمال النماذج الثلاث والجدول رقم(3) يوضح الترتيب الذي اتبع في التطبيق.

جدول رقم(3)

ترتيب تطبيق نماذج الاختبار على المجموعات الثلاث

الترتيب	المجموعة		
	الأولى	الثانية	الثالثة
أولا	النموذج الأول	النموذج الثالث	النموذج الثاني
ثانيا	النموذج الثاني	النموذج الأول	النموذج الثالث
ثالثا	النموذج الثالث	النموذج الثاني	النموذج الأول

وقد بلغ عدد [illegible] الأشكال الثلاثة من الاختبار التحصيلي (211) طالبا وطالبة، بعد تغيب عشرة من أصل عينة الدراسة التجريبية، والتي كانت عددها (221)، كما وجد بعد تصحيح الاختبار ان هناك(14) ورقة كانت إجابة المفحوصين عليها غير جادة، فأسقطت هذه الورقات، ليصبح عدد أفراد عينة الدراسة التجريبية، التي خضعت نتائجهم للتحليل الإحصائي(197) طالبا وطالبة.

تصحيح النماذج الثلاثة من الاختبار التحصيلي:

1- النموذج الأول من الاختبار : حسبت علامة المفحوص الكلية على فقرات هذا النموذج، باستخدام معادلة التصحيح لأثر التخمين(العقاب)، حيث تم جمع الإجابات الصحيحة، وكذلك الإجابات غير الصحيحة التي حاولها كل مفحوص، أما الفقرات المحذوفة فقد اعتبرت فقرات لم يحاولها المفحوص، وبذلك أصبحت علامة المفحوص =

$$\text{عدد الإجابات الصحيحة} - \frac{\text{عدد الإجابات غير الصحيحة التي حاولها المفحوص}}{\text{عدد البدائل} - 1}$$

وتتضمن تعليمات هذا النموذج ما يشير إلى تجنب التخمين كما يلي:

أجب عن السؤال إذا استطعت استبعاد بديلاً أو اكثر من البدائل المعروضة، أو كــان لديــك حــدس بان أحد البدائل هو الإجابة الصحيحة. أما إذا لم تســتطع اســتبعاد أيٍ من البدائل، ولم يكن لديك حدس فالأفضل ألا تجيب علــى السؤال، لأن حصولك على أي إجابة خاطئة سوف يعرضك للعقاب من خلال خصم جزء من العلامات التي تحصلُها في الاختبار.

2- الــنموذج الثانــي من الاختبار: حسبت علامة المفحوص الكلية على فقــرات هــذا الــنموذج، باســتخدام معادلــة التصــحيح لأثــر التخمين(المكافأة)، حيث تم جمع الإجابات الصحيحة، وكذلك الفقرات المحذوفــة التــي لم يحاول المفحوص الإجابة عليها، وبذلك أصبحت علامة المفحوص =

$$\text{عدد الإجابات الصحيحة} + \frac{\text{عدد الإجابات المحذوفة}}{\text{عدد البدائل}}$$

وتتضمن تعليمات هذا النموذج ما يشير إلى تجنب التخمين كما يلي :

أجب عن السؤال إذا استطعت استبعاد بديلاً أو اكثر من البدائل المعروضة، أو كان لديك حدس بان أحد البدائل هو الإجابة الصحيحة. أما إذا لم تستطع اســتبعاد أي بديــل، ولم يكن لديك حدس فالأفضل ألا تجيب على السؤال، وستكافأ على الفقرات المتروكة دون إجابة.

3- الــنموذج الثالــث من الاختبار: حسبت علامة المفحوص الكلية على فقرات هذا النموذج، باستخدام معادلة التصحيح لأثر التخمين(المكافأة والعقاب)، حيث تم جمع الإجابات الصحيحة بالإضافة إلى الإجابات غير الصحيحة التي حاولها كل مفحوص، و الفقرات المحذوفة التي لم يحاول المفحوص الإجابة عليها، وبذلك أصبحت علامة المفحوص =

$$\text{عدد الإجابات الصحيحة} - \frac{\text{عدد الإجابات غير الصحيحة}}{\text{عدد البدائل} - 1} + \frac{\text{عدد الإجابات المحذوفة}}{\text{عدد البدائل}}$$

وتتضمن تعليمات هذا النموذج ما يشير إلى تجنب التخمين كما يلي :

أجب عن السؤال إذا استطعت استبعاد بديلاً أو اكثر من البدائل المعروضة، أو كان لديك حدس بان أحد البدائل هو الإجابة الصحيحة. أما إذا لم تستطع اســتبعاد أي بديــل، ولم يكن لديك حدس فالأفضل ألا تجيب على السؤال. لأنك سوف تعاقب على الإجابات الخاطئة وتكافئ على الإجابات المحذوفة.

المعالجات الإحصائية:

هدفت هذه الدراسة إلى المقارنة بين أثر ثلاث طرق لضبط التخمين على كل من معامل ثبات الاختبار، والخطأ المعياري للقياس، وشكل توزيع درجات الاختبار ولذا تم إجراء المعالجات التالية:

1-اختــبار الفــروق بين معاملات ثبات الأداء على النماذج الثلاثة للاختــبار، باســتخدام الإحصائي (t) لاختبار الفروق بين معاملات الارتباط للعينات المستقلة، عند مستوى الدلالة الإحصائية ($\alpha \geq 05$).

$$t = (\rho_{12} - \rho_{13})\sqrt{\frac{(N-3)(1+\rho_{23})}{2(1-\rho_{12}^2-\rho_{13}^2-\rho_{23}^2+2\rho_{12}\rho_{13}\rho_{23})}}$$ (عودة و الخليلي،2000)

2- اختبار الفروق بين القيمة المقدرة للخطأ المعياري للقياس، باستخدام الإحصائي t للفرق بين تباينين غير مستقلين.

$$t = \frac{\sigma_1^2 - \sigma_2^2}{\sqrt{\frac{4\sigma_1^2\sigma_2^2}{N-2}(1-\rho^2)}}$$ (عودة و الخليلي،2000)

3- وصــف الفرق في شكل توزيع الأداء على النماذج الثلاثة للاختبار في ضوء خصائص التوزيع في كل منها.

الفصل الرابع

عرض النتائج

يتـــناول هذا الفصل عرضا لنتائج التحليل الإحصائي الهادفة للإجابة عن أسئلة الدراسة.

أولا: النتائج المتعلقة بالسؤال الأول والثاني:

1 هــل لطريقة التصحيح اثر على ثبات الاختبار؟ أي هل توجد فروق ذات دلالـــة إحصـــائية بيــن معاملات ثبات الاختبار عند تصحيحه بالطرق الثلاث موضوع الدراسة ؟

2- كــيف تؤثــر طريقة التصحيح على الخطأ المعياري للقياس؟ أي هل توجد فروق ذات دلالة إحصائية بين قيمة الخطأ المعياري للقياس لكل طريقة من طرق التصحيح الثلاث؟

وللإجابــة عن السؤال الأول والثاني تم حساب معامل الثبات والخطأ المعــياري للقياس لعلامات جميع المفحوصين في عينة الدراسة التجريبية، وعلى كل نموذج من النماذج الثلاثة للاختبار التحصيلي، والتي تم تصحيح كــل نموذج منها بإحدى طرق التصحيح الثلاثة والجدول رقم(4) يبن تلك النتائج.

يشير رقم النموذج في هذا الحدول إلى نموذج الاختبار والطريقة التي تــم بهــا تصحيح تلك النموذج على النحو التالي:(1) تعني النموذج الأول والطــريقة الأولى لتصحيحه، باستخدام معادلة العقاب على التخمين، (2) تعنــي النموذج الثاني والطريقة الثانية لتصحيحه، باستخدام معادلة المكافأة علــى عدم التخمين، (3) تعني النموذج الثلث والطريقة الثلثة لتصحيحه، باستخدام معادلة الجمع بين المكافأة والعقاب.

جدول رقم(4)

معاملات الثبات و قيم الخطأ المعياري للقياس لعلامات جميع المفحوصين على النماذج الثلاثة للاختبار.

النموذج	عدد المفحوصين	عدد الأسئلة	معامل الثبات	الخطأ المعياري للقياس
1	197	30	0.81	2.75
2	197	30	0.84	2.31
3	197	30	0.86	2.26

ويظهر من هذا الجدول، أنه عند استخدام طريقة المكافأة و العقاب لتصحيح الاختبار، كان معامل الثبات أعلى منه عند استخدام طريقة معادلة التصحيح لأثر التخمين(العقاب) وطريقة المكافأة على عدم التخمين.

وللإجابة عن السؤال الأول، اختبرت الفروق بين معملات ثبات النماذج الثلاثة للاختبار، باستخدام الإحصائي(ت) لاختبار الفروق بين معاملات الارتباط للعينات المرتبطة "Correlated Samples".

والجدول رقم(5) [illegible] معاملات الثبات، للنماذج الثلاثة من الاختبار.

الجدول رقم(5)

قيم(t) المحسوبة للفروق بين معاملات الثبات عند عينة الدراسة.

الشكل	الوسط الحسابي	الانحراف المعياري	معامل الثبات	المقارنات الثنائية لمعاملات الثبات: نموذج	قيمة (t)
1	9.8	5.79	0.81	1و2	1.22
2	15.08	6.06	0.84	2و3	1.52
3	11.79	6.31	0.86	1و3	*2.22

* ذات دلالة إحصائية ($\alpha \le .05$)

يتضح من الجدول رقم(5)، أن الفرق بين معامل ثبات الاختبار المصحح باستخدام معادلة العقاب،ومعامل ثبات الاختبار المصحح باستخدام

معادلة المكافأة، لم يصل لمستوى الدلالة الإحصائية($\alpha \leq 05$)هذا على الرغم من أن معامل ثبات الاختبار المصحح باستخدام معادلة المكافأة أكبر من معامل ثبات الاختبار المصحح باستخدام معادلة العقاب. كذلك تشير النتائج إلي أن الفرق بين معامل ثبات الاختبار المصحح باستحدام معادلة المكافأة، ومعامل ثبات الاختبار المصحح باستخدام معادلة المكافأة والعقاب، لم يكن دالاً مع أن معامل ثبات الاختبار المصحح باستخدام معادلة المكافأة والعقاب أكبر من معامل ثبات الاختبار عندما يصحح باستخدام معادلة المكافأة. ويشير الجدول إلى أن الفرق بين معامل ثبات الاختبار المصحح باستخدام المكافأة والعقاب، ومعامل ثبات الاختبار المصحح باستخدام معادلة العقاب، كان دالاً احصائياً عند مستوى($\alpha \leq 05$) ، وكان الفرق لصالح استحدام معادلة المكافأة والعقاب.

ان هذه النتائج تمكن من الإجابة عن السؤال الأول بالقول بأن لطريقة التصحيح لأثر التخمين اثر على معامل ثبات الاختبار، وان معامل ثبات الاختبار المصحح باستخدام معادلة المكافأة والعقاب يكون أعلى من معامل ثبات الاختبار المصحح باستخدام معادلة العقاب أو معادلة المكافأة، في حين لا يوجد فرق بين معامل ثبات الاختبار المصحح باستخدام معادلة العقاب، ومعامل ثبات الاختبار المصحح باستخدام معادلة المكافأة.

للإجابة عن السؤال الثاني، اختبرت الفروق بين قيم الخطأ المعياري للقياس للنماذج الثلاثة للاختبار، باستخدام الإحصائي t للفرق بين تباينين غير مستقلين والجدول رقم(6) يبين، قيم(t) المحسوبة للفروق بين قيم الخطأ المعياري للقياس، للنماذج الثلاثة من الاختبار.

الجدول رقم(6)

قيم(t) المحسوبة للفروق بين قيم الخطأ المعياري للقياس عند عينة الدراسة.

الشكل	الوسط الحسابي	الانحراف المعياري	الخطأ المعياري للقياس	المقارنات الثنائية لقيم الخطأ المعياري للقياس	
				نموذج	قيمة (t)
1	9.8	5.79	2.75	1 و 2	3.05*
2	15.08	6.06	2.31	2 و 3	0.55
3	11.79	6.31	2.26	3 و 1	4.17*

* ذات دلالة إحصائية ($\alpha \geq .05$).

يتضح من الجدول رقم(6)، ان الفرق بين قيمة الخطأ المعياري للقياس المحسوبة للنموذج الأول، وقيمة الخطأ المعياري للقياس المحسوبة للنموذج الثاني، كان ذو دلالة إحصائية، وقد كان هذا الفرق لصالح استخدام معادلة المكافأة. وكذلك فان الفرق بين قيمة الخطأ المعياري للقياس المحسوبة للنموذج الأول وقيمة الخطأ المعياري للقياس المحسوبة للنموذج الثالث، كان أيضا ذو دلالة إحصائية، لصالح استخدام معادلة المكافأة والعقاب. أما الفرق بين قيمة الخطأ المعياري المحسوبة للنموذج الثاني والثالث فانه لم يكن دالاً وان كان يشير إلى ان قيمة الخطأ المعياري للقياس تكون اقل باستخدام معادلة المكافأة والعقاب.

ان هذه النتائج تمكننا من الإجابة عن السؤال الثاني المتعلق بأثر لطريقة التصحيح لأثر التخمين اثر على قيمة الخطأ المعياري للقياس وان قيمة هذا الخطأ تكون اقل باستخدام معادلة التصحيح المعتمد على المكافأة والعقاب.

ثانياً: النتائج المتعلقة بالسؤال الثالث:

هــل تختلف خصائص توزيع درجات الاختبار باختلاف طريقة التصحيح لأثر التخمين ؟

وللإجابــة عن هذا السؤال، تم استخراج مقاييس النزعة المركزية ومقاييس التشـــتت لعلامـــات المفحوصـــين على الأشكال الثلاثة للاختبار كما هي موضحه في الجدول رقم(7).

جدول رقم(7)

مقايس النزعة المركزية والتشتت لعلامات المفحوصين على النماذج الثلاثة للاختبار

النموذج	مقاييس النزعة المركزية			مقاييس التشتت		
	الوسط الحسابي	[illegible]	[illegible]	[illegible]	التباين	المدى
1	9 8	[illegible]	[illegible]	[illegible]	33 49	27
2	15 08	[illegible]	[illegible]	[illegible]	36 67	27
3	11 79	11	11	6 31	39 83	28

ويتضح من الجدول رقم(7)، ان أعلى قيمة للوسط الحسابي والوسيط كانت لعلامات المفحوصين على النموذج الثاني المصحح باستخدام معادلة المكافــأة، وأقــل قيمة له كانت لعلامات المفحوصين على النموذج الأول المصحح باستخدام معادلة العقاب. أما المنوال فيتضح من الجدول ان القيمة التـــي تركزت عندها علامات النموذج الثاني والثالث متساوية وأعلى من القيمة التي تركزت عندها علامات النموذج الأول. ويظهر الجدول أيضا ان أعلى قيمة للانحراف المعياري كانت لعلامات المفحوصين على نموذج الثالـــث، وأقــل قيمة كانت لعلامات المفحوصين على النموذج الأول، أما المدى فقد كان متساوياً للنماذج الثلاثة تقريباً.

مقاييس النزعة المركزية أو الموضع(Measures of Location)

أولاً: الوسط الحسابي(Mean)

يتضح من الجدول السابق(7) ان الوسط الحسابي لعلامات المفحوصين في النموذج الثاني المصحح باستخدام معادلة المكافأة كان أعلى من الوسط الحسابي في النموذج الأول والثالث. ولاختبار الفروق بين هذه الأوساط تم استخدام الإحصائي(t) لاختبار الفروق بين الأوساط الحسابية للعينات المرتبطة.والجدول رقم(8) يبن قيم(t) المحسوبة عند مستوى دلالة $(\alpha \leq 05)$ للفروق بين الأوساط الحسابية لعلامات المفحوصين على النماذج الثلاثة.

جدول رقم(8)

المقارنات [illegible] علامات [illegible] المفحوصين على النماذج الثلاثة.

النموذج	الوسط الحسابي	المقارنة بين الأوساط الحسابية: الشكل	المقارنة بين الأوساط الحسابية: قيمة(t)
1	9.8	2 و 1	*14.12
2	15.08	3 و 2	*13.23
3	11.79	3 و 1	*6.55

*ذات دلالة إحصائية $(\alpha \leq .05)$

يتضح من الجدول: وجود فروق ذات دلالة إحصائية عند مستوى $(\alpha > 05)$ بين الأوساط الحسابية لعلامات المفحوصين على النموذج الأول والثاني، و لصالح النموذج الثاني. وتظهر النتائج أيضا وجود فروق ذات دلالة إحصائية عند مستوى الدلالة$(\alpha \leq 05)$ بين الأوساط الحسابية لعلامات المفحوصين على النموذج الثاني والثالث، و لصالح النموذج الثاني. كذلك فقد بينت النتائج وجود فروق ذات دلالة إحصائية بين الأوساط الحسابية

لعلامات المفحوصين على النموذج الأول والثالث، حيث كانت تلك الفروق لصالح النموذج الثالث.

وهـــذا يعنـــي ان هـــناك اثـــر لطـــريقة التصحيح على متوسط أداء المفحوصـــين، فالأوســـاط الحســـابية لعلامات المفحوصين على الاختبار المصـــحح باســـتخدام معادلـــة المكافأة، كانت أعلى من الأوساط الحسابية لعلامـــات المفحوصـــين عند تصحيح الاختبار باستخدام معادلة العقاب أو استخدام معادلة العقاب والمكافأة، عند مستوى الدلالة الإحصائية ($\alpha \leq 05$).

ثانيا: الوسيط(median)

يتضـــح من الجدول رقم(7) ان قيمة الوسيط كانت أعلى عند تصحيح الاختـــبار باســـتخدام معادلـــة المكافأة وبمستوى(14) وهذا يعني ان نصف المفحوصين كانت علاماتهم أقل من أو تساوي 14 درجة والنصف الأخر اكـــبر مـــن أو يســـاوي(14) درجـــة. وهذه العلامة قريبة من علامة القطع الـــنظرية(15). في حين بلغت قيمة الوسيط(11) عندما تم تصحيح الأداء بـــناءاً على معادلة المكافأة والعقاب، بينما بلغت قيمة الوسيط عند تصحيح الاختبار باستخدام معادلة العقاب(8).

مقاييس التشتت:

الانحراف المعياري والتباين

يتضـــح مـــن الجدول رقم(7) ان أقل قيمة للانحراف المعياري كانت لـــتوزيع علامـــات النموذج الأول من الاختبار المصحح باستخدام معادلة العقـــاب ممـــا يدل على ان علامات هذا التوزيع أكثر تجانساً من علامات الـــنماذج الأخـــرى، بينما نلاحظ ان أكبر قيمة للانحراف المعياري كانت

لتوزيع علامات النموذج الثالث المصحح باستخدام معادلة المكافأة و العقاب، مما يدل ان علامات هذا التوزيع أقل تجانساً.

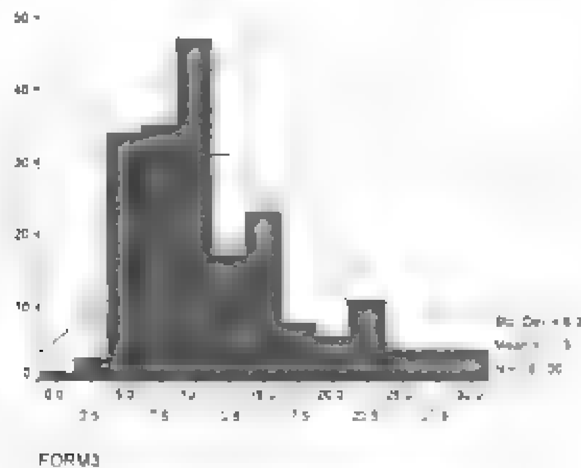
وصف توزيع علامات الاختبار:

من خلال استعراض مقاييس النزعة المركزية، وأشكال المدرج التكراري لتوزيع علامات النماذج الثلاثة للاختبار الموضحة بالأشكال رقم(5)،(6)،(7)،نجد ان جميع هذه التوزيعات موجبة الالتواء "Positively Skewed"ولكن بدرجات مختلفة، كما ان هذه التوزيعات حادة "Leptokurtic" ولكن بدرجات تفلطح مختلفة، والجدول رقم(9) يوضح ذلك.



شكل(6) المدرج التكراري لعلامات النموذج الثاني

شكل(5) المدرج التكراري لعلامات النموذج الأول



شكل(7) المدرج التكراري لعلامات النموذج الثالث

جدول رقم(9)

نوع الالتواء والتفلطح ودرجة كل منهما لتوزيع علامات المفحوصين على النماذج الثلاثة للاختبار

النموذج	الالتواء	درجة الالتواء	التفلطح	درجة التفلطح
1	موجب	0.16	حاد	0.20
2	موجب	0.20	حاد	0.21
3	موجب	0.14	حاد	0 24

يتضـــح من الجدول رقم(9) ومن خلال مقارنة خصائص التوزيعات الــثلاثة لعلامـــات المفحوصـــين على النماذج الثلاثة للاختبار بخصائص التوزيع الطبيعي، نجد أن توزيع علامات المفحوصين على النموذج الثالث اقــرب مـــا يكون للتوزيع الطبيعي، فدرجة التواءه اقرب إلى الصفر الدال علـــى درجة التواء التوزيع الطبيعي، ودرجة التفلطح له اقرب إلى 0.26 التي تدل على درجة التفلطح للتوزيع الطبيعي(عودة، [illegible] خليلي،2000).

الفصل الخامس

مناقشة النتائج والتوصيات

هدفت هذه الدراسة إلى المقارنة بين ثلاثة طرق لتصحيح اختبارات الاختيار من متعدد لأثر التخمين على كل من : معامل ثبات الاختبار، الخطأ المعياري للقياس، شكل توزيع الاختبار.

أولاً: مناقشة النتائج المتعلقة بالسؤال الأول والثاني:

1- هل لطريقة التصحيح اثر على ثبات الاختبار ؟

2- كيف تؤثر طريقة التصحيح على الخطأ المعياري للقياس؟

أظهرت نتائج التحليل الإحصائي للفروق بين معاملات الثبات، وجود فروق بين معاملات الثبات، تعزى لطريقة التصحيح لأثر التخمين. فقد كانت هذه الفروق، ذات دلالة إحصائية عند مستوى الدلالة($\alpha \geq 05$) لصالح استخدام معادلة التصحيح لأثر التخمين(المكافأة والعقاب)، مقارنة بمعادلة التصحيح لأثر التخمين(العقاب) في حين لا يوجد فرق بين معامل ثبات الاختبار المصحح باستخدام معادلة العقاب، ومعامل ثبات الاختبار المصحح باستخدام معادلة المكافأة.

كذلك أظهرت نتائج التحليل الإحصائي، للفروق بين قيم الخطأ المعياري للقياس، وجود فروق بين قيم الخطأ المعياري للقياس، تعزى لطريقة التصحيح لأثر التخمين. فقد كانت هذه الفروق، ذات دلالة إحصائية عند مستوى الدلالة($\alpha \geq 05$) لصالح استخدام معادلة التصحيح لأثر التخمين(المكافأة والعقاب)، مقارنة بمعادلة التصحيح لأثر التخمين(العقاب) وكذلك كانت ذات دلالة إحصائية عند مستوى الدلالة($\alpha \geq 05$) لصالح

استخدام معادلة التصحيح لأثر التخمين(المكافأة والعقاب)، مقارنة بمعادلة التصحيح لأثر التخمين (المكافأة). في حين لا يوجد فرق بين معامل ثبات الاختـبار المصـحح باسـتخدام معادلـة العقاب، ومعامل ثبات الاختبار المصحح باستخدام معادلة المكافأة.

ان تفسـير هذه النتائج، يمكن ان يكون في ضوء الاستراتيجيات التي يتـبعها المفحوص عند اخذ الاختبار، والتعليمات المرافقة لكل من الطرق الـثلاث لتصحيح أثر التخمين، وأثر ذلك على تباين علامات المفحوصين علـى الاختـبار. فعـندما يعطـى المفحوص التعليمات المرافقة لطريقة التصـحيح الأولـى التـي تركز على معاقبة المفحوص إذا قام بالتخمين سيكون هناك حد من قبل المفحوص للتخمين، ولكن ذلك يعتمد على طبيعة المفحـوص، فطبـيعة المفحوص تدفعه دائما للحصول على علامة عالية، فالمفحوص تحت تعليمات المعادلة الأولى سيجيب على الفقرات التي يعرف إجابـتها وتبقى الفقرات الأخرى التي لا يعرف المفحوص إجابتها أو لديه معرفة جزئية عن هذه الفقرات، فهو أمام حالتين، الأولى أن يترك الفقرات دون إجابـة ويلتزم بتعليمات معادلة التصحيح وستبقى علامته هي مجموع الفقـرات التـي يعرف إجابتها. أو أن لا يلتزم بتعليمات المعادلة ويحاول تخميـن بعـض هذه الفقرات على أمل أن يكون تخمينه صحيحاً. وبما ان العقـاب يولـد أحيانا التمرد على التعليمات فان بعض المفحوصين سيعمل على تخمين بعض الفقرات، وبالتالي يزداد التباين الكلي في الأخطاء ويقل معامل الثبات وهذا يؤدي إلى زيادة قيمة الخطأ المعياري للقياس .

فـي حيـن ان تعليمات المعادلة الثانية تعطي المفحوص شعوراً بأنه سيكافأ على الفقرات المتروكة دون إجابة(التي لا يعرف إجابتها) وبالتالي سيزيد ذلك من العلامة الكلية له، سيدفع به للابتعاد عن التخمين مقارنة مع

تعليمات المعادلة الأولى، وبالتالي سيقل التباين الكلي في الأخطاء مما يؤدي إلى زيادة معامل الثبات ونقصان في قيمة الخطأ المعياري للقياس، ولكن بعض المفحوصين سيضحي بتخمين بعض الفقرات وخصوصا إذا فكر بالطريقة التالية " إذا جاءت الإجابة صحيحة على الفقرة سأحصل على علامة وان جاءت خاطئة فلن اخسر شيئاً ".أما تحت تعليمات المعادلة الثالثة فإن العقاب والمكافأة يحتمعان وبالتالي لن يفكر المفحوص بالتفكير السابق وسيلتزم إلى حد كبير بالتعليمات ولن يلجأ للتخمين، وبالتالي يقل تباين الأخطاء ويزداد معامل الثبات.

ثالثاً : مناقشة النتائج المتعلقة بالسؤال الثالث:

هل تختلف خصائص توزيع درجات الاختبار باختلاف طريقة التصحيح لأثر التخمين ؟

أظهرت نتائج التحليل الإحصائي للفروق بين الأوساط الحسابية لعلامات المفحوصين وجود فروق ذات دلالة إحصائية عند مستوى دلالة ($\alpha \geq 05$) تعزى لطريقة التصحيح لأثر التخمين، ولصالح طريقة التصحيح القائمة على المكافأة.

ويمكن تفسير هذه النتائج، في ضوء خصائص كل معادلة من معادلات التصحيح. فإذا كان هناك اختبار يتكون من 10 فقرات اختيار من متعدد(ذات بدائل أربع)، وكان أحد المفحوصين يعرف إجابة 6 فقرات، و تقدم لهذا الامتحان تحت تعليمات المعادلة الأولى(معادلة العقاب) والتزم بتعليماتها فانه بتطبيق المعادلة على إجاباته ستكون علامته 6، وان لم يلتزم وخمن الفقرات التي لا يعرف إجابتها فانه حسب الاحتمال النظري للتخمين العشوائي سيحصل على علامة واحدة زيادة على علامته الحقيقية ولكن

اجـــراءات هذه المعادلة سوف تعاقبه على ذلك بخصم علامة وبالتالي فإن علامته ستبقى كما هي ان لم تنقص.

إما إذا تقدم للامتحان تحت تعليمات المعادلة الثالثة والتزم بتعليماتها فانه بتطبيق المعادلة على إجاباته فستكون علامته 7، وان لم يلتزم وخمن الفقـــرات التي لا يعرف إجابتها فلن يؤدي ذلك لزيادة علامته وربما يؤدي ذلك إلى نقصانها.

أمـــا إذا تقدم للامتحان تحت تعليمات المعادلة الثانية والتزم بتعليماتها فانه بتطبيق المعادلة على إجاباته فستكون علامته 7، وان لم يلتزم وخمن الفقرات التي لا يعرف إجابته، حسب الاحتمال النظري للتخمين العشوائي سيحصـــل على علامة واحدة زيادة على علامته الحقيقية وبالتالي ستزداد علامته ويكون تحصيله أفضل من تعرضه للتخمين.

أظهـــرت النـــتائج أيضـــا اختلاف شكل توزيع العلامات تبعاً لطريقة التصحيح لأثر التخمين فقد كان شكل توزيع علامات النموذج الثالث اقرب إلى التوزيع الطبيعي.

ويمكـــن تفســـير هذه النتيجة، في ضوء تعريف العلامة الحقيقية التي تمـــثل درجة امتلاك الفرد للخاصية أو السمة بدون أخطاء، فمن المفترض ان الأفـــراد يـــتوزعون توزيعاً طبيعياً على متصل السمة، وبما ان طريقة التصـــحيح الثالـــثة تعطـــي اقل قيمة للخطأ فهي بذلك تعطي افضل تقدير لعلامة الفرد الحقيقية، وبالتالي سيكون توزيع هذه العلامات اقرب ما يكون للتوزيع الطبيعي.

التوصيات

في ضوء النتائج السابقة فإنّ الدراسة الحالية توصي بما يلي:

استـــخدام معادلـــة التصـــحيح لأثر التخمين(العقاب والمكافأة) في تصحيح اختبارات الاختيار من متعدد. وذلك لانها تعطي اكبر تقدير لمعامل الثبات.



قائمة المراجع

أ- المراجع العربية :

أبـو ريـنـة، فريد. (1992). أساسيات القياس والتقويم في التربية، ط1، مكتبة الفلاح، الكويت.

أبـو لبدة، سبع. (1985). مبادئ القياس النفسي و التقييم التربوي،عمان، الأردن.

الثبيتي، علي. (1998). صدق البناء النظري لأدوات جمع المعلومات في البحوث التربوية والنفسية وانعكاس ذلك على تفسير النتائج، رسالة الخليج العربي، العدد 69، السنة 19، ص 17-62.

الحسـين، أحمـد. (2000). [illegible] في الإنترنت للمعلمين والمعلمات. http //www gecoites com teacher club exam htm

الخطيـب، سـلطان. (1993). القياس والتقويم التربوي في المدرسة، مؤسسة شيرين، عمان، الأردن.

سواقد، ساري. (1983). المقارنة بين أثر ثلاث طرق لتصحيح الاختبار مـن مـتعدد علـى الخصـائص السيكومترية للاختبار وعلى أداء المفحوصـين من مستويات تحصيل ودرجات مخاطرة مختلفة على الاختـبار، رسـالة ماجستير غير منشورة، جامعة اليرموك، اربد، الأردن.

سـواقد، سـاري. (1992). اختـبار صحة الافتراضات النظرية لطرق التصـحيح لأثـر التخميـن ومقارنة اثر استخدام هذه الطرق على الخصائص السيكومترية للفقرة،رسالة دكتوراه غير منشورة، الجامعة الأردنية، عمان، الأردن.

الشرقاوي، أنور؛ الشيخ، سليمان؛كاظم، أمينة؛وعبد السلام، نادية.(1996). اتجاهــات معاصــرة فــي القياس والتقويم النفسي والتربوي، مكتبة الأنجلو المصرية، القاهرة.

عباس، محمد. (1993). المقارنة بين خمس طرق لتقدير الخطأ المعياري الشرطي في القياس عند مستويات محددة لعلامات الاختبار، رسالة دكتوراه غير منشورة، الجامعة الأردنية، عمان، الأردن.

عبيدات، ذوقان؛عباس، عبد الرحمن؛ وعبد الحق، كايد. (1982). البحث العلمـــي، مفهومه، أدواته، أساليبه، دار مجدلاوي للنشر والتوزيع، عمان، الأردن .

علام، صلاح الدين. (2002). القياس والتقويم التربوي والنفسي أساسياته وتطبيقاته وتوجيهاته، دار الفكر العربي، القاهرة.

عـــلام، صلاح الدين. (1990). أثر المواقف الاختبارية جماعية المرجع ومحكية المرجع في مستويات التخمين وسلوك المخاطرة والأداء في علم النفس، مجلة علم النفس، العدد(17)، المجلد الخامس ، ص98-107.

عــودة، أحمــد؛ الخليلـــي، خليل. (2000). الإحصاء للباحث في التربية والعلوم الإنسانية، ط2، دار الأمل للنشر والتوزيع،اربد، الأردن.

عودة، احمد. (2000). القياس والتقويم في العملية التدريسية، عمان.

عــودة، محمــد فضل موسى. (1993). أثر بنية فقرة الاختيار من متعدد علــى الخصـــائص السيكومترية لكل من الفقرات والاختبار، رسالة دكتوراه غير منشورة، الجامعة الأردنية، عمان، الأردن.

ملحـم، سـامي. (2000). القياس والتقويم في التربية وعلم النفس، ط1، دار المسيرة للنشر والتوزيع والطباعة، عمان، الأردن.

ب- المراجع الأجنبية

Abu-Sayf, F. (1977) . A new formula score. Educational and psychological Measurement, Vol.37, PP. 853-862.

Alnabhan, Mousa. (2002) .An empirical investigation of the effects of three methods of handling guessing and risk taking on the psychometric indices of a test, Social Behavior And Personality, Vol. 30, NO.(7), PP. 645 – 65[illegible].

Ben-shaker, G. & Sinai, Y. (1991). Gender differences in multiple-choice: the role of differential guessing tendencies, Journal of educational measurement, Vol. 28,NO.(1), PP.23-35

Ben-simon,A.& Budescu ,D.& Nevo, B.(1997). A comparative study of measures of partial knowledge in multiple choice tests, Applied psychological measurement, Vol. 21,No.(1), PP.65-88.

Bliss, L.(1980).A test of Lord's assumption regarding examinee Guessing behavior on multiple choice tests using elementary school students, Journal of educational measurement, Vol.17,No.(2), PP.147-153

Brown, F. (1983). Principles of educational and psychological testing,. (3rd ed), Holt Rinehart and Winston, Inc, New York.

Budescu,D.& Bar-Hillel,M.(1993).To guess or not to guess:A decision- theoretic view of formula scoring, Journal of educational measurement, Vol. 30, NO .(4) ,PP.277-291.

Bush , m.(2001). A Multiple Choice Test that Rewards Partial Knowledge, Journal of further and higher education, Vol. 25,No.(2).

Crocker,L.& Algina, J.(1986), Introduction to classical and modern test theory,University of Florida,U.S.A..

Cronbach, L. (1984). Essentials of psychological testing, Harper and Row, U.S.A.

Cross, L. Frary, R (1[illegible]) An empirical test of Lord's theoretical results regarding formula scoring of multiple-choice tests. Journal of educational measuremant. Vol. 14, PP.313-321 .

EBEL,R.(1979) Essentials of Educational Measurement, 3rd edn(Englewood Cliffs, New Jersey, Prentice-Hall).

Frary ,R.(1989). Partial-credit scoring methods for multiple-choice tests, Applied measurement In education, Vol. 2,No.(1), PP.79-96.

Gibbons, J & Olkin, I. & Sobel, M.(1979). A subset selection technique for scoring items on a multiple choicetest, Psychometrika. Vol. 44,No.(3), PP.259-270

Grounlund ,Norman E.(1985). Measurement and evaluation in teaching,. five edition ,University of Illinois,U.S.A.

Ineke,H & Wilfried,A.& Paul,M. (2002). Scores on a yes no vocabulary test:correction for guessing and response style, *<u>Language Testing</u>* Vol.19,NO.(3), PP. 227–245

Jaradat, D.& Tollefson,M.(1988). The impact of alternative scoring procedures for multiple-choice items on test reliability, validity, and grading, <u>Educational and psychological measurement</u>, Vol. 48, PP.627-635.

Jaradat,D. & Sawagad,S.(1986).The subset selection technique for multiple choice tests: An Empirical Inquary, <u>Journal of educational measuremant</u>,Vol 23, NO.(4), PP.369-376

Meherns, William A. & Lehmann, Irvin J.(1978). <u>Measurement And Evaluation in Education And Psychology</u>,(2nd ed), Holt Rinehart Winston, USA.

Muijtjens, A.&Vanmameren, H.& Hoogenboom, R. & Evers,J. & Vnder,V. (1999). The effect of a 'don't know' option on test scores: number-right and formula scoring compared, <u>Medical education</u>, Vol. 33NO. (4), pp. 267-275.

Nunnally, J.C.(1978). <u>Psychometric Theory,</u> (2nd ed), New York. Mccrow-Hill.

Thorndike, R.L. & Hagen, E.P.(1977). <u>Measurement and evaluation in psychology and education</u>, 4th,ed.New york, Wiley.

ملحق رقم (1)

الشكل الأولي للاختبار التحصيلي

ملحق رقم(1)

1 أي العبارات التالية تصف بدرجة أكثر دقة وشمولا العلاقة بين القياس والتقويم ؟

أ- التقويم يستند إلى القياس ويعتمد على نتائجه.

ب التقويم يشوبه عنصر الذاتية بينما القياس موضوعي.

ج- القياس يستمد مباشرة من التقويم.

د- القياس متطلب أساسي للتقويم.

2- أي أنواع الاختبارات الآتية يمكن استخدامها بدرجة أفضل في التقويم الختامي لمقرر دراسي معين ؟

أ- اختبار تشخيصي عقب الانتهاء من كل وحدة يشتمل عليها المقرر

ب- اختبار تحصيلي مقنن مرتبط بالمقرر.

ج- اختبار مدرسي عادي.

د- اختبار استعداد دراسي.

3- أي مما يأتي يعد تعريفا أفضل للاختبار النفسي أو التربوي ؟

[illegible]

ب- [illegible] ات الفرد

ج- مجموعة من المثيرات التي صممت بحيث تستدعي استجابات معينة

د- مجموعة من [illegible] التي يمكن عن طريقها جمع ملاحظات عن سمات

[illegible] الأفراد.

4- تصنف الاختبارات إلى اختبارات مرجعية المعيار، واختبارات مرجعية المحك استناداً إلى ؟

أ طريقة تفسير نتائجها

ب نوع مفرداتها

ج مدى قياسها لتحصيل الطلاب

د- ما إذا كانت مقننة أم من تصميم المعلم

5- الميزة الرئيسية التي تتميز بها الاختبارات المقننة عن الاختبارات غير المقننة هي ؟

أ- الصدق

ب- الموضوعية

ج- توافر المعايير

د- جودة فقراتها.

6- الفرق الأساسي بين اختبارات الاستعدادات واختبارات التحصيل يتعلق :

أ- بالهدف من استخدام الاختبار.

ب- بطبيعة أو نوع القدرة المقاسة.

ج- بأسلوب القياس.

د بمدى ضيق أو اتساع محتوى الاختبار.

7 أي من أدوات القياس التالية يتطلب تطبيقها وتفسير نتائجها معلومات متخصصة أكثر من غيرها

أ- اختبارات الذكاء الجمعية.

ب- الاختبارات الإسقاطية للشخصية

ج- مقاييس الشخصية التي تعتمد على التقارير الذاتية للمفحوصين

د- بطاريات الاختبارات التحصيلية.

8 إذا لم يلتزم المعلم بتعليمات الاختبارات المقننة عند تطبيقها وتصحيحها التزاما تاما، فمن المحتمل أن

أ- تصبح معايير الاختبار غير مناسبة.

ب- يتساهل المعلم في مراقبة الطلاب.

ج- يعطي المعلم فرصا غير متكافئة للطلاب في الإجابة

د- يتهاون الطلاب أثناء إجابتهم عن الاختبار.

9- أي مما يأتي يعبر بدرجة أفضل عن مفهوم "معايير" الاختبارات المقننة ؟

أ- مستويات ينبغي أن يحققها المفحوصين من ذوي الأداء المتوسط

ب- مستويات أداء ينبغي أن يحققها معظم المفحوصين

ج- درجات محولة تعتمد على أداء معظم المفحوصين في عينة التقنين

د- درجات محولة تمثل الدرجات الخام التي يحصل عليها المفحوصين في عينة التقنين

10 في أي مما يأتي نتوقع أن يبحث معلم عن "معايير" اختبار معين ؟

أ- البحوث التي استخدمت الاختبار

ب- الدوريات السيكولوجية والتربوية

ج- دليل الاختبار

د- كل من أ و ب.

11- ما الخطوة الأخيرة التي تتبع عادة في بناء الاختبارات التحصيلية المقننة ؟

أ- المراجعة النهائية لفقرات الاختبار و[illegible]

ب- تطبيق الاختبار على عينة كبيرة ممثلة من الطلاب

ج- تقييم محتوى فقرات الاختبار بواسطة مجموعة من [illegible]

د- التحليل الإحصائي لفقرات الاختبار.

12- إذا أراد معلم أن يتأكد مما إذا كان اختبار تحصيلي معين يقيس الأهداف [illegible] معين، فمن الأفضل أن يقوم بفحص :

أ- فقرات الاختبار

ب- المراجعات الناقدة للاختبار

ج- دليل الاختبار.

د- الدراسات الحديثة التي استخدمت الاختبار.

13 يعتمد تفسير نتائج الاختبارات التحصيلية المعدة على تحديد :

أ- النسبة المئوية من الأهداف التي أتقنها الطلاب.

ب- النسبة المئوية من الفقرات التي أجاب عنها الطالب إجابة صحيحة

ج- طبيعة ومستوى الفقرات التي لم يستطع الطالب الإجابة عنها.

د- مركز الطالب بالنسبة لزملائه الذين أعد الاختبار من أجلهم.

14 عندما يطلق مفهوم الموضوعية على اختبار تربوي أو نفسي فانه يعني :

أ- طريقة تنظيم فقرات الاختبار

ب- دقة تقنين الاختبار.

ج- اتساق عملية تصحيح الاختبار.

د- مدى تمثيل فقرات الاختبار للمحتوى الذي تقيسه.

15- يقصد بمفهوم "التقنين" في القياس التربوي و النفسي :

أ- المعيار الذي ينسب إليه أداء الفرد.

ب- مستوى الأداء الذي ينبغي أن يحققه الفرد متوسط القدرة.

ج- توحيد ظروف تطبيق أدوات القياس وإجراءاتها.

16- أي مما يأتي ينبغي



ج- لا يكون غاية في الصعوبة

د- أن يكون منشورا لدى إحدى دور النشر.

17- أي مما يأتي يعبر بدرجة أكثر دقة عما تقيسه اختبارات الذكاء ؟

أ- عينة مقننة من سلوك الفرد.

ب- القدرة الفطرية لدى الفرد.

ج- مستوى نضج الفرد.

د- النجاح المستقبلي المحتمل للفرد.

18- تهدف اختبارات الذكاء لقياس الأداء الأقصى، بينما تهدف مقاييس الشخصية لتقدير :

أ- السلوك المعتاد المميز للفرد ز

ب- المحددات اللاشعورية لسلوك الفرد.

ج- مدى تأثر الفرد بعمليات القياس.

د- أداء الفرد في مواقف واقعية.

19- أي مما يأتي يصف بدرجة أفضل اختبار بينيه للذكاء ؟

أ- يقيس ذكاء الأفراد مهما اختلفت خلفياتهم الثقافية قياسا عادلا.

ب- يقيس القدرة الفطرية قياسا عادلا.

ج- يقيس القدرة العقلية غير متأثرة بمتغيرات الشخصية الأخرى.

د- محتواه مشبع بالجانب اللفظي.

20- يمكن الإفادة من مقاييس الميول المقننة بدرجة أفضل في :

أ- التوجيه التربوي والمهني.

ب- تخطيط المناهج الدراسية.

ج- تقديم تعليم علاجي للطلاب.

د- تأكيد مفهوم الذات لدى الطلاب.

21- أي مما يأتي يعد عيبا رئيساً من عيوب الأنواع المختلفة لمقاييس الشخصية ؟

أ- يستغرق تطبيقها وقتا كبيراً.

ب- تتطلب توافر محك موضوعي لتقدير صدقها.

ج- يمكن للمفحوصين تزوير إجاباتهم بسهولة

د- [illegible] المفحوصين

22- قيمة وسط مجموعة [illegible]

[illegible]

د- 26.5

23- قيمة الانحراف المعياري لمجموعة الدرجات 2، 4، 6، 8، 10 تساوي تقريبا :

أ- 1.4

ب- 2.0

ج- 2.8

د- 3.2

24- مقياس النزعة المركزية الذي يصف بدرجة أفضل التوزيعات التي تشتمل على درجات متطرفة هو

أ- المنوال.

ب- الوسيط.

ج- المتوسط.

د- كل من أ و ب .

25- توزيعا الدرجات اللذان يمثلهما الشكل الموضح يختلفان في :

أ- الالتواء فقط.

ب- التشتت فقط.

ج- النزعة المركزية فقط.

د- التشتت والنزعة المركزية

26 تحويل الدرجات الخام إلى درجات معيارية أو درجات تائية يتطلب الحصول على :

أ- الوسيط.

ب- الانحراف المعياري.

ج- المدى المطلق.

د- السؤال.

27- طُبق اختبار على مجموعة من طلاب نهاية المرحلة الثانوية، ووجد أن معامل الارتباط بين درجاتهم في الاختبار، ودرجاتهم السنة الأولى بعد التحاقهم بالجامعة 0,82 فيمكننا القول بأن :

أ- الاختبار منبئ جيد للنجاح في الدراسة الجامعية.

ب- درجات الاختبار غير متسقة.

ج- حوالي 82% من الطلاب سوف ينجحون في دراساتهم الجامعية.

د- الطلاب الذين حصلوا على درجات أقل من 82 يحتمل أن يخفقوا في دراستهم الجامعية

28 القيمة المرتفعة لمعامل الارتباط بين اختبارين أحدهما يقيس الاستعداد الدراسي والآخر يقيس الذكاء يدل على أن الاختبارين يقيسان :

مصدرين مختلفين من مصادر السلوك

[illegible]

[illegible]

[illegible]

29- يعد المنحنى الطبيعي في القياس النفسي والتربوي لأنه

أ- يعطي تفسيرا دقيقا لدرجات الاختبار

ب- يمثل بدرجة كبيرة شكل كثير من السمات الإنسانية

ج- يعتمد على الأساليب الإحصائية.

د- يسهل تطبيقه وفهم نتائجه.

30 إذا طُبق اختبار على 1000 طالب وكانت درجاته تتوزع توزيعا طبيعيا، فما عدد الطلاب الذين يحتمل أن تنحصر درجاتهم بين +1، −1 انحراف معياري عن المتوسط :

أ- 160

ب- 340

ج- 500

د- 680

31 يذكر طالب أنه حصل على درجة الخام 68 في اختبار العلوم أي نوع من المعلومات يمكن أن يساعد بدرجة أفضل في فهم معنى هذه الدرجة فيما يتعلق بتحصيل الطالب في العلوم ؟

أ- اشتمال الاختبار على 90 سؤال.

ب- نصف طلاب الفصل رسبوا في الاختبار.

ج- متوسط درجات الفصل 65

د- أعلى درجة في الفصل 83

32- أجـــاب طالب عن 25 سؤال إجابة صحيحة في اختبار رياضيات يشتمل 50 سؤال. ما التفسير المحتمل لأداء الطالب ؟

أ- يحتاج إلى دروس تقوية في الرياضيات

ب- يعرف حوالي نصف مادة الرياضيات التي تضمنها الاختبار.

ج- مركز الطالب هو المئين 50 .

د- المعلومات المعطاة غير كافية لتفسير.

33 متوســـط درجـــات اختبار تحصيلي مقنن 50، وانحرافه المعياري 10، فإذا حصل طالب على الدرجة 65 في الاختبار، فانه يمكننا الحكم على أداء هذا الطالب بأنه :

أ- متميز.

ب- يفوق المتوسط بقليل

ج- متوسط.

د- يقترب من المتوسط.

34- إذا حصـــل طالب على الدرجة الخام 35 في اختبار معين، ومتوسط درجات الصفة 37.5 وانحرافها 8.4 فإن درجته المعيارية تساوي تقريبا :



35- في اختبار [illegible] المئين المناظر للدرجة الخام 35 =

أ- 2

ب- 16

ج- 50

د- 84

36- إذا حصل طالب في اختبار على درجة خام تناظر المئين 56. فهذا يعني أن :

أ- أداءه يفوق 56 طالبا من أقرانه من المفحوصين.

ب- أداء 56% من الطلاب أفضل من أدائه في الاختبار.

ج- أداءه يعادل أو أفضل من أداء 56% من أقرانه من المفحوصين.

د- درجة الطالب تزيد بقدر 6 درجات عن متوسط درجات الاختبار.

37- فـيما يلي المتوسط و الانحراف المعياري لدرجات اختبارين على الترتيب 40، 5 34، 2 فإذا كانت هذه الدرجـــات تتوزع توزيع طبيعي، وحصل طالب على الدرجتين 45، 37 في كل من الاختبارين على الترتيب فأي مما يأتي يعد صحيحا فيما يتعلق بأداء الطالب ؟

أ- أداؤه في الاختبار الأول أفضل من أداؤه في الاختبار الثاني

ب- أداؤه في الاختبار الثاني أفضل من أداؤه في الاختبار الأول.

ج- أداؤه متعادل في كل من الاختبارين.

د- المعلومات المعطاة غير كافية للحكم على أداء الطالب.

38- إذا كان العمر الزمني لطفل 4 سنوات و 6 شهور، وعمره العقلي كما يقاس باختبار بينية 6 سنوات و 3 شهور، فإن نسبة ذكائه تساوي :

أ- 72
ب- 100
ج- 128
د- 152

39- أي أنواع الصدق الآتية يناسب بدرجة أكبر الاختبارات التحصيلية ؟

أ- صدق التكوين الفرضي
ب- الصدق التلازمي.
ج- الصدق التنبؤي.
د- صدق المحتوى.

40- يمكن تقدير صدق اختبار نفسي بإيجاد الارتباط بين درجات :

أ- الفقرات الزوجية والفردية التي تشتمل عليها الاختبار.
ب- صورتين متكافئتين من الاختبار.
ج- الاختبار ودرجات محك معين مستقل عنه.
د- تطبيقين متتاليين للاختبار نفسه.

41- أي الخصائص الآتية تتأثر بدرجة اكبر بقلة عدد فقرات الاختبار ؟

أ- الموضوعية.
ب- التقنين.
ج- الثبات.
د- الصدق.

42- أي الطـرق الآتية المستخدمة في تقدير ثبات درجات الاختبارات تأخذ بعين الاعتبار أخطاء التباين التي ترجع إلى اختلاف كل عينات المفحوصين وظروف تطبيق الاختبار ؟

أ- الاتساق الداخلي
ب- التجزئة النصفية.
ج- الصور المتكافئة
د- إعادة إجراء الاختبار.

43- إذا كان معامل الارتباط بين النصفين الزوجي والفردي لفقرات اختبار 0,74 فإن القيمة التقديرية لمعامل الثبات بعد تصحيحه بمعادلة سبيرمان – براون تساوي :

أ- 0 80
ب- 0 85
ج 0 90
د- 0 95

44- يحتمل أن تنخفض قيمة معامل ثبات الاختبار إذا

أ- زادت فقراته

ب- تشتت درجاته تشتتا كبيرا

ج- زادت صعوبة فقراته بالنسبة لمعظم المفحوصين.

د- اشتمل على فقرات لا تقيس ما وضع لقياسه.

45- طبق اختبار مرتين في يومين متتاليين على مجموعة من الطلاب، ووجد أن معامل الارتباط بين الدرجات في المرتين 0,95 أي الاستنتاجات الآتية المتعلقة بالدرجات صحيحة ؟

أ- الدرجات تتميز بدرجة عالية من الثبات.

ب- الدرجات تتميز بدرجة عالية من الصدق

ج- الدرجات غير مستقرة إلى حد كبير.

د- الدرجات غير مميزة.

46- تتميز الاختبارات التي يكتبها المعلم للطلاب عن الاختبارات التحصيلية المقننة من حيث أنها

أ- تقدم أدلة أفضل عما استوعبه التلاميذ من الحقائق والمعارف.



التي يود المعلم تحقيقها لدى تلاميذه.

التلاميذ الأقوياء والتلاميذ الضعاف

47- أي الأهداف التالية [illegible]

أ- تنمية الحقائق المعرفية

ب- تفسير أسباب حدوث الظواهر الجغرافية

ج- تذكر بعض المصطلحات الجغرافية

د- [illegible] خرائط [illegible]

48- المستوى المعرفي للهدف التعليمي ترجم الطالب قصة نثرية من اللغة العربية إلى اللغة الإنجليزية هو

أ- التذكر

ب- [illegible]

ج- [illegible]

د- التحليل

49- أي أنواع الفقرات الآتية تتأثر إجابات الطلاب عنها بدرجة أكبر بعامل التخمين ؟

أ- التكملة.

ب- الصواب والخطأ

ج- الاختيار من متعدد.

د- المزاوجة.

50 جميع ما يأتي من مميزات فقرات الاختيار من متعدد فيما عدا أن هذه الفقرات

أ- يمكن أن تستخدم في قياس الفهم والتطبيق.

ب- تستخدم بكثرة في مختلف مجالات المعرفة.

ج- تتطلب من المفحوص استخدام عمليات التعرف

د- تحتاج إجاباتها إلى قدر كبير من المهارات اللغوية.

51 أهم ما تتميز بها الاختبارات الموضوعية عن اختبارات المقال :

أ- توفير وقت المعلم.

ب- قياسها مستويات مختلفة من الأهداف التعليمية.

ج- اشتمالها على عينة أكثر تمثيلا لمحتوى المادة الدراسية.

[illegible]لة الوقت اللازم للإجابة عنها.

52- [illegible] معلم [illegible] فإنه يفضل استخدام سؤال :

أ- [illegible]

ب- [illegible]

ج- [illegible]

د- [illegible]

53- [illegible] جيدة لأنها

[illegible]عية في السهولة.

ب- مختصرة جدا

ج- تقيس تذكر الحقائق.

د- تحتمل أكثر من إجابة واحدة

54 يرى خبراء القياس أن الأسئلة الاختيارية في اختبارات المقال :

أ- مطلوبة لأنها توفر الحرية لكل [illegible]

ب- مطلوبة لأنها تجعل الاختبار ممثلا [illegible] الدراسي

ج- غير مرغوبة لأنها تجعل الموازنة [illegible] درجاتهم في الاختبار غير ممكنة

د- غير مرغوبة لأنها تضيع وقت [illegible] سيجيبون عنها

55- يمكن أن يتبع المعلم جميع الطرق الآتية عند تقدير درجات أسئلة اختبارات المقال فيما عدا

أ- استقطاع بعض الدرجات للأخطاء النحوية وما شابه ذلك.

ب- الإعداد المسبق لنموذج تصحيح مناسب.

ج- استبعاد أسماء الطلاب من أوراق الإجابة قبل بدء التصحيح.

د- تصحيح كل سؤال في جميع أوراق الإجابة قبل الانتقال إلى السؤال التالي.

56 يقترح كثير من خبراء القياس عدم استخدام صيغة تصحيح أثر التخمين في فقرات الاختيار من متعدد، والاستعاضة عن ذلك :

أ- بزيادة عدد فقرات الاختبار.

ب- بإعطـاء تعليمات للطلاب باللجوء إلى التخمين المبني على ما لديهم من معلومات في حالة عدم تأكدهم من الإجابة الصحيحة.

ج- بتقليل الزمن المطلوب للإجابة.

د- بإعطاء تعليمات للطلاب بعدم اللجوء إلى التخمين على الإطلاق.

57- يفيد "جدول مواصفات الاختبار" بدرجة أساسية في :

أ- كتابة أسئلة واضحة وجيدة

ب- انتقاء نوع الفقرات المناسبة لقياس الأهداف التعليمية.

ج- تمثيل فقرات الاختبار للأهداف التعليمية المتعلقة بالمحتوى الدراسي.

د- تنظيم محتوى الاختبار

58- في الاختبارات التي يكتبها المعلم لتلاميذه ينبغي ترتيب الفقرات



[illegible] الصواب والخطأ وانتهاء بأسئلة المقال.

[illegible] المقال وانتهاء بفقرات الصواب والخطأ .

[illegible] المقال وانتهاء بفقرات الصواب والخطأ .

59 [illegible] طبق اختبار على 40 طالبا [illegible] إحدى فقراته

فإن معامل صعوبة هذه الفقرة يساوي

أ- 0.06

ب- 0.15

ج- 0.34

د- 0.85

60 اختبر 55 طالبا باختبار معين، وحصل 15 طالبا على أعلى الدرجات، 15 طالبا على أدنى الدرجات. فإذا أجاب 10 طلاب في المجموعة العليا، [illegible] طلاب في المجموعة الدنيا عن إحدى فقرات الاختبار إجابة صحيحة، فإن معامل تميز هذه الفقرة يساوي :

أ- صفر

ب- 0,33

ج- 0,50

د- 1,00

ملحق رقم (2)

الشكل النهائي للاختبار التحصيلي

ملحق رقم(2)

1- أي العبارات التالية تصف بدرجة أكثر دقة وشمولا العلاقة بين القياس والتقويم ؟

أ- التقويم يستند إلى القياس ويعتمد على نتائجه.

ب- التقويم قد يتأثر بالذاتية بينما القياس موضوعي.

ج- القياس يستمد مباشرة من التقويم.

د- القياس متطلب أساسي للتقويم

2- أي مما يأتي يعد تعريفا أفضل للاختبار النفسي أو التربوي ؟

أ- مقياس مباشر لعينة من السلوك

ب- مقياس غير مباشر لمجموعه معارف وقدرات لدى الفرد

ج- مجموعة من المثيرات التي تصمم بحيث تستدعي استجابات معينة

د- مجموعة من المطالب المقننة التي يمكن عن طريقها جمع ملاحظات عن سمات أو خصائص الأفراد

3- تصف الاختبار [illegible] المحك تستند الى ؟

[illegible]

4 الميزة الرئيسية التي تتميز بها الاختبارات المقننة عن الاختبارات غير المقننة هي ؟

أ الصدق

ب- الموضوعية

ج- توافر المعايير

د- جودة فقراتها

5 من عيوب الأسئلة المقالية بشكل عام.

أ درجة التخمين فيه مرتفعة

ب لا تقيس قدرات عقليه عليا

ج- غير صادقة وثابتة

د- فرصة الغش فيها مرتفعة

6- عندما يطلق مفهوم الموضوعية على اختبار تربوي أو نفسي فانه يعني :

أ- طريقة تنظيم فقرات الاختبار

ب- دقة تقنين الاختبار

ج- اتساق عملية تصحيح الاختبار.

د- مدى تمثيل فقرات الاختبار للمحتوى الذي تقيسه.

7- إذا طبق اختبار على مجموعة من طلاب نهاية المرحلة الثانوية، ووجد أن معامل الارتباط بين درجاتهم في الاختبار، ودرجاتهم في السنة الأولى بعد التحاقهم بالجامعة 0.82 فيمكن القول بأن

أ- الاختبار متنبئ جيد للنجاح في الدراسة الجامعية.

ب- درجات الاختبار غير متسقة.

ج- حوالي 82% من الطلاب سوف ينجحون في دراساتهم الجامعية.

د- الطلاب الذين حصلوا على درجات أقل من 82 يحتمل أن يخفقوا في دراستهم الجامعية .

8- يفضل استخدام المنحنى الطبيعي في القياس النفسي والتربوي لأنه :

أ- يضمن تفسيرا دقيقا لدرجات الاختبار.

ب- يماثل بدرجة كبيرة شكل توزيع كثير من السمات الإنسانية.

ج- يعتمد على الأساليب الإحصائية.

د- يسهل تطبيقه وفهم نتائجه.

9- جاب طالب عن 25 سؤالا إجابة صحيحة في اختبار [illegible] يشتمل على 50 سؤال [illegible] التفسير المحتمل لأداء الط[illegible]

[illegible]

[illegible] على نصفها الاختبار

[illegible]

د- المعلومات المعطاة غير كافية للتفسير.

10 متوسط درجات اختبار تحصيلي مقن 50، وانحرافه المعياري 10، فإذا حصل طالب على الدرجة 65 في الاختبار، فإنه يمكننا الحكم على أداء هذا الطالب بأنه :

أ- متميز.

ب- يفوق المتوسط بقليل.

ج- متوسط.

د- يقترب من المتوسط.

11 أي أنواع الصدق الآتية يناسب بدرجة أكبر الاختبارات التحصيلية ؟

أ- صدق التكوين الفرضي

ب- الصدق التلازمي.

ج- الصدق التنبؤي.

د- صدق المحتوى.

12- يمكن تقدير صدق اختبار نفسي بإيجاد الارتباط بين درجات :

الفقرات الزوجية والفردية التي اشتمل عليها الاختبار.

ب- صورتين متكافئتين من الاختبار.

ج- نتائج الاختبار ودرجات محك معين مستقل عنه.

د- تطبيق الاختبار نفسه مرتين متتاليتين.

13. أي الخصائص الآتية تتأثر بدرجة أكبر بقلة عدد فقرات الاختبار ؟

أ- الموضوعية.

ب- التقنين

ج- الثبات

د- الصدق

14 أي الطرق الآتية المستخدمة في تقدير ثبات درجات الاختبارات تأخذ بعين الاعتبار أخطاء التباين التي ترجع إلى اختلاف كل عينات المفحوصين وظروف تطبيق الاختبار

أ- الاتساق الداخلي.

ب- التجزئة النصفية.

ج- الصور المتكافئة.

د- إعادة إجراء الاختبار.

15 إذا كان معامل الارتباط بين النصفين الزوجي والفردي لفقرات اختبار 0.74 فإن القيمة التقديرية لمعامل الثبات بعد تصحيحه بمعادلة سبيرمان – براون تساوي :

[illegible]

-16 يحتمل أن تنحصر [illegible]

ب- تشتت درجاته تشتتا كبيرا.

ج- زيادة صعوبة فقراته بالنسبة لمعظم المفحوصين

د- اشتمل على فقرات لا تقيس ما وضع لقياسه

-17 طبق اختبار مرتين في يومين متتاليين على مجموعة من الطلاب، ووجد أن معامل الارتباط بين الدرجات في المرتين 0.95 أي الاستنتاجات الآتية المتعلقة بالدرجات صحيحة ؟

أ- الدرجات تتميز بدرجة عالية من الثبات.

ب- الدرجات تتميز بدرجة عالية من الصدق.

ج- الدرجات غير مستقرة إلى حد كبير.

د- الدرجات غير مميزة.

18 أي الأهداف التعليمية الآتية يعد هدفا سلوكيا ؟

أ- أن يفهم الطالب الحقائق الجغرافية

ب- أن يفسر الطالب أسباب حدوث الظواهر الجغرافية

ج- أن يتذكر الطالب بعض المصطلحات الجغرافية

د- أن يألف الطالب خرائط الطقس

19 اكثر اشكال الأسئلة الموضوعية تأثراً بالتخمين هي أسئلة :

أ- التكملة

ب- الصواب والخطأ

ج- الاختيار من متعدد

د- المزاوجة

20- جميع ما يأتي هي من مميزات فقرات الاختيار من متعدد عدا أن هذه الفقرات :

أ- يمكن أن تستخدم في قياس الفهم والتطبيق.

ب- تستخدم بكثرة في مختلف مجالات المعرفة.

ج- تتطلب من المفحوص استخدام عمليات التعرف.

د- تحتاج إجاباتها إلى قدر كبير من المهارات اللغوية.

21- أهم ما تتميز بها الاختبارات الموضوعية عن اختبارات المقال :

أ- توفير وقت المعلم

ب- قياسها مستويات مختلفة من الأهداف التعليمية.

ج- اشتمالها على عينة اكثر تمثيلا لمحتوى المادة الدراسية.

د- قلة الوقت اللازم للإجابة عنها.

22- [illegible] معلم لاجتماعيات التحقق مما إذا كان تلاميذه يستطيعون التعرف على عواصم خمس دول، [illegible]

فانه يفضل استخدام سؤال :

أ- مقال.

ب- تكملة

ج- صواب أو خطأ

د- اختيار من متعدد

23- "طوكيو أهم مدينة في اليابان" تعد فقرة صواب أو خطأ غير جيدة [illegible]

أ- غاية في السهولة.

ب- مختصرة جدا.

ج- تقيس تذكر الحقائق.

د- تحتمل أكثر من إجابة واحدة.

24 يمكن أن يتبع المعلم جميع الطرق الآتية عند تقدير درجات أسئلة المقال [illegible]

أ- خصم بعض الدرجات للأخطاء النحوية [illegible]

ب- الإعداد المسبق لنموذج تصحيح مناسب.

ج- استبعاد أسماء الطلاب من أوراق الإجابة قبل بدء التصحيح.

د- تصحيح كل سؤال في جميع أوراق الإجابة قبل الانتقال إلى السؤال التالي

25- يقترح كثير من خبراء القياس عدم استخدام صيغة التصحيح لأثر التخمين في فقرات الاختيار من متعدد والاستعاضة عن ذلك :

أ- بزيادة عدد فقرات الاختبار

ب- بإعطاء تعليمات للطلاب باللجوء إلى التخمين المبني على ما لديهم من معلومات في حالة عدم تأكدهم من الإجابة الصحيحة.

ج- بتقليل الزمن المطلوب للإجابة

د- بإعطاء تعليمات للطلاب بعدم اللجوء إلى التخمين على الإطلاق.

26- يفيد "جدول مواصفات الاختبار" بدرجة أساسية في :

أ- كتابة أسئلة واضحة وجيدة.

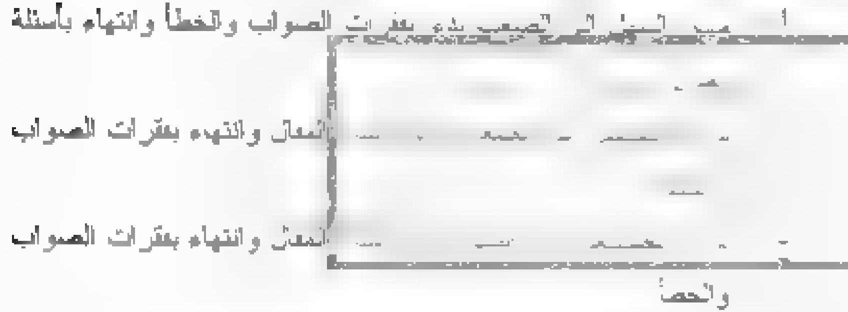
ب- انتقاء نوع الفقرات المناسبة لقياس الأهداف التعليمية.

ج- تمثيل فقرات الاختبار للأهداف التعليمية المتعلقة بالمحتوى الدراسي.

د- تصميم محتوى الاختبار

27- في الاختبارات التي يكتبها المعلم لتلاميذه ينبغي ترتيب الفقرات :

أ- من السهل إلى الصعب بدءاً بفقرات الصواب والخطأ وانتهاء بأسئلة



ب- ... المقال وانتهاء بفقرات الصواب

ج- ... المقال وانتهاء بفقرات الصواب والخطأ

د- بطريقة عشوائية

28- إذا طبق اختبار على 40 طالباً، أجاب 6 طلاب إجابة خطأ عن إحدى فقراته، فإن معامل صعوبة هذه الفقرة يساوي

أ- 0.06
ب- 0.15
ج- 0.34
د- 0.85

29- اختبر 55 طالباً باختبار معين، وحصل 15 طالب على أعلى الدرجات و 15 طالب على أدنى الدرجات. فإذا أجاب 10 طلاب في المجموعة العليا و 10 طلاب في المجموعة الدنيا عن إحدى فقرات الاختبار إجابة صحيحة، فإن معامل تمييز هذه الفقرة يساوي :

أ- صفر
ب- 0.33
ج- 0.50
د- 1.00

30- من عيوب الاسئلة الموضوعية ذات البديلين (صح خطأ) انها

أ- فرصة التخمين فيها مرتفعة

ب- تقيس قدرات عقلية عب

ج- تتسب المستويات الب

د- تغطي محتوى المادة



ع

ملحق رقم (3)

الفقرات المحذوفة

ملحق رقم(3)

رقـم الفقرة المحذوفة	الفقــرة

-2 أي أنواع الاختبارات الآتية يمكن استخدامها بدرجة أفضل في التقويم الختامي لمقرر دراسي معين ؟

أ- اختبار تشخيصي عقب الانتهاء من كل وحدة يشتمل عليها المقرر

ب- اختبار تحصيلي مقن مرتبط بالمقرر.

ج- اختبار مدرسي عادي.

د- اختبار استعداد دراسي

6 الفرق الأساسي بين اختبارات الاستعدادات واختبارات التحصيل يتعلق

أ- بالهدف من استخدام الاختبار

ب- بطبيعة أو نوع القدرة المقاسة.

ج- بأسلوب القياس.

[illegible]

-7 أي من أدوات القياس [illegible] تخصصه أكثر من غيرها

[illegible]

[illegible]

ج- [illegible] الذاتية للمفحوصين

د- بطاريات الاختبارات التحصيلية

8 إذا لم يلتزم المعلم بتعليمات الاختبارات المقننة عند تطبيقها وتصحيحها التزاما تامة، فمن المحتمل أن :

أ تصبح معايير الاختبار غير مناسبة.

ب- يتساهل المعلم في مراقبة الطلاب.

ج- يعطى المعلم فرصا غير متكافئة للطلاب في الإجابة

د- يتهاون الطلاب أثناء إجاباتهم عن الاختبار.

-9 أي مما يأتي يعبر بدرجة أفضل عن مفهوم " معايير " الاختبارات المقننة ؟

أ مستويات ينبغي أن يحققها المفحوصين من ذوى الأداء المتوسط

ب- مستويات أداء ينبغي أن يحققها معظم المفحوصين

ج درجات محولة تعتمد على أداء معظم المفحوصين في عينة التقنين

د- درجــات محولة تناظر الدرجات الخام التي يحصل عليها المفحوصين في عينة

10 في أي مما يأتي يتوقع أن يبحث معلم عن " معايير " اختبار مقن ؟

أ البحوث التي استخدمت الاختبار.

ب الدوريات السيكولوجية والتربوية

ج دليل الاختبار.

د- كل من أ و ب.

11 ما الخطوة الأخيرة التي تتبع عادة في بناء الاختبارات التحصيلية المقننة ؟

أ- المراجعة النهائية لفقرات الاختبار وتعليمات تطبيقه.

ب- تطبيق الاختبار على عينة كبيرة ممثلة من الطلاب.

ج- تقيم محتوى فقرات الاختبار بواسطة مجموعة من الخبراء

د- التحليل الإحصائي لفقرات الاختبار.

12 [illegible] معلم أن يتأكد مما [illegible] كان اختبار تحصيلي مقنن يقيس أهداف محتوى درسي معين، فمن الأفضل أن يقوم بفحص :

أ- فقرات الاختبار.

ب- المراجعات الناقدة للاختبار.

ج- دليل الاختبار.

د- الدراسات الحديثة التي استخدمت الاختبار.

-13 يعتمد تفسير نتائج الاختبارات التحصيلية المقننة على تحديد :

أ- النسبة المئوية من الأهداف التي أتقنها الطلاب.

[illegible] عليها الطالب إجابة صحيحة.

[illegible] الطالب للإجابة عنه

[illegible] الاختبار من أجلهم

-15 [illegible] بمفهوم [illegible]

[illegible] المعيار الذي ينسب إليه أداء الفرد.

ب- مستوى الأداء الذي ينبغي أن يحققه الفرد متوسط القدرة

ج- توحيد ظروف تطبيق أدوات القياس وإجراءاتها.

د- اتساق عملية تصحيح أدوات القياس.

16 أي مما يأتي ينبغي مراعاته بدرجة أكبر عند انتقاء اختبار مقنن ؟

أ- أن يكون الاختبار صادقا.

ب- أن تكون إجراءات تطبيقه غير معقدة.

ج- ألا يكون غاية في الصعوبة.

د- أن يكون منشورا لدى إحدى دور النشر.

-17 أي مما يأتي يعبر بدرجة أكثر دقة عما تقيسه اختبارات الذكاء ؟

أ- عينة مقننة من سلوك الفرد.

ب- القدرة الفطرية لدى الفرد.

ج- مستوى نضج الفرد.

د- النجاح المستقبلي المحتمل للفرد.

18- تهدف اختبارات الذكاء لقياس الأداء الأقصى، بينما تهدف مقاييس الشخصية لتقدير :

أ- السلوك المعتاد المميز للفرد ز

ب- المحددات اللاشعورية لسلوك الفرد.

ج- مدى تأثر الفرد بعمليات القياس.

د- أداء الفرد في مواقف واقعية.

19- أي مما يأتي يصف بدرجة أفضل اختبار بينيه للذكاء ؟

أ- يقيس ذكاء الأفراد مهما اختلفت خلفياتهم الثقافية قياسا عادلا.

ب- يقيس القدرة الفطرية قياسا عادلا.

ج- يقيس القدرة العقلية غير متأثرة بمتغيرات الشخصية الأخرى.

د- محتواه مشبع بالجانب اللفظي.

20- يمكن الإفادة من مقاييس الميول المقننة بدرجة أفضل في :

أ- التوجيه التربوي والمهني.

ب- تخطيط المناهج الدراسية.

ج- تقديم تعليم علاجي للطلاب.

د- تأكيد مفهوم الذات لدى الطلبة.

21- أي مما يأتي يعد عيبا رئيسا من عيوب الأنواع المختلفة لمقاييس الشخصية ؟

أ- يستغرق تطبيقها وقتا كبيراً.

ب- تتطلب توافر محك موضوعي لتقدير صدقها.

ج- يمكن للمفحوصين تزوير إجاباتهم بسهولة.

د- يصعب إعداد خطة محددة لتصحيح إجابات المفحوصين.

22- قيمة وسيط مجموعة الدرجات 27، 50، 13، 5، 46، 34، 63 تساوي :

أ- 25.0
ب- 34.0
ج- 35.5
د- 26.5

23- قيمة الانحراف المعياري لمجموعة الدرجات 2، 4، 6، 8، 10 تساوي تقريبا :

أ- 1.4
ب- 2.0
ج- 2.8
د- 3.2

24- مقياس النزعة المركزية الذي يناسب بدرجة أفضل التوزيعات التي تشتمل على درجات متطرفة هو :

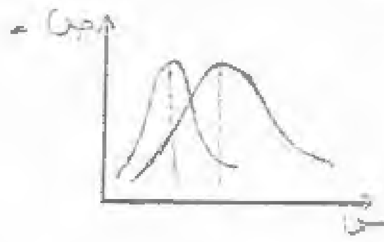
أ- المنوال.

ب- الوسيط.

ج- المتوسط.

د- كل من أ و ب .

25- توزيعا الدرجات اللذان يمثلهما الشكل الموضح يختلفان في :



أ- الالتواء فقط.

ب- التشتت فقط.

ج- النزعة المركزية فقط.

د- التشتت والنزعة المركزية

26- تحويل الدرجات الخام إلى درجات معيارية أو درجات تائية يتطلب الحصول على :

أ- الوسيط.

ب- الانحراف المعياري.

ج- المدى المطلق.

د- المنوال.

28- القيمة المرتفعة لمعامل الارتباط بين اختبارين أحدهما يقيس الاستعداد الدراسي والآخر يقيس الذكاء يدل على أن الاختبارين يقيسان :

أ- مظهران مختلفان من مظاهر السلوك.

ب- قدرات موروثة لدى الأفراد.

ج- جانب واحد من جوانب السلوك.

د- التحصيل وليس القدرة العقلية.

30- إذا طـــبق اختبار على 1000 طالب وكانت درجاته تتوزع توزيعا طبيعيا، فان عدد الطلاب الذين يحتمل أن تنحصر درجاتهم بين +1، -1 انحراف معياري عن المتوسط :

أ- 160

ب- 340

ج- 500

د- 680

31- ذكر طالب أنة حصل على درجة الخام 68 في اختبار العلوم. أي نوع من المعلومات يمكن أن يساعد بدرجه أفضل في فهم معنى هذه الدرجة فيما يتعلق بتحصيل الطالب في العلوم ؟

أ- اشتمال الاختبار على 90 سؤال.

ب- نصف طلاب الفصل رسبوا في الاختبار.

ج- متوسط درجات الفصل 65.

د- أعلى درجة في الفصل 83.

34- إذا حصـــل طالب على الدرجة الخام 35 في اختبار معين، ومتوسط درجات فصلة 37,5 وانحرافها 8,4 فان درجته المعيارية تساوي تقريبا :

أ- 0,44-

ب- 0,30-

ج- 0,23 +

د- 0,27 +

35- إذا أعطـــي اختبار متوسط درجاته الخام 50 وانحرافها المعياري 15، فان المئين المناظر للدرجة الخام 35 يساوي :

أ- 2
ب- 16
ج- 50
د- 84

36- إذا حصل طالب في اختبار على درجة خام تناظر المئين 56. فهذا يعني أن :

أ- أداءه يفوق 56 طالبا من أقرانه من المفحوصين.
ب- أداء 56% من الطلاب أفضل من أدائه في الاختبار.
ج- أداءه يعادل أو أفضل من أداء 56% من أقرانه من المفحوصين.
د- درجة الطالب تزيد بقدر 6 درجات عن متوسط درجات الاختبار.

37- فيما يلي المتوسط والانحراف المعياري لدرجات اختبارين على الترتيب 40، 5 . 34، 2 . فإذا كانت هذه الدرجات تتوزع توزيع طبيعي، وحصل طالب على الدرجتين 45، 37 في كل من الاختبارين على الترتيب. فأي مما يأتي يعد صحيحا فيما يتعلق بأداء الطالب ؟

أ- أداؤه في الاختبار الأول أفضل من أداؤه في الاختبار الثاني.
ب- أداؤه في الاختبار الثاني أفضل من أداؤه في الاختبار الأول.
ج- أداؤه متعادل في كل من الاختبارين.
د- المعلومات المعطاة غير كافية للحكم على أداء الطالب.

38- إذا كـــان العمـــر الزمنـــي لطفل 4 سنوات و 6 شهور، وعمره العقلي كما يقاس باختبار بينية 6 سنوات و 3 شهور، فان نسبة ذكائه تساوي :

أ- 72
ب- 100
ج- 128
د- 152

46- تتميز الاختبارات الصفية التي يكتبها المعلم لتلاميذ عن الاختبارات التحصيلية المقننة من حيث أنها :

أ- تقدم أدلة أفضل عما استوعبه التلاميذ من الحقائق والمعارف.
ب- أيسر في تطبيقها وتصحيحها.
ج- أكثر صدقا في قياس الأهداف التعليمية التي يود المعلم تحقيقها لدى تلاميذه.
د- أكثر دقة في قياس الفروق الفردية بين التلاميذ الأقوياء والتلاميذ الضعاف.

48- المستوى المعرفي للهدف التعليمي " يترجم الطالب قطعة نثرية من اللغة العربية الى اللغة الإنجليزية " هو.

أ- التذكر.
ب- الفهم.
ج- التطبيق.
د- التحليل.

54- يرى خبراء القياس أن الأسئلة الاختيارية في اختبارات المقال :

أ- مطلوبة لأنها توفر العدالة لكل طالب.

ب- مطلوبة لأنها تجعل الاختبار ممثلا بدرجة أفضل للمحتوى الدراسي.

ج- غير مرغوبة لأنها تجعل الموازنة بين الطلاب على أساس درجاتهم في الاختبار غير ممكنة

د- غير مرغوبة لأنها تضيع وقت الطلاب في انتقاء الأسئلة التي سيجيبون عنها.

